DOCUMENT RESUME

ED 306 224 TM 012 916

AUTHOR Zimowski, Michele F.; Wothke, Werner
TITLE The Measurement of Human Variation in Spatial Visualizing Ability: A Process-Oriented Perspective. Technical Report No. 1986-1.
INSTITUTION Johnson O'Connor Research Foundation, Chicago, IL. Human Engineering Lab.
PUB DATE Sep 86
NOTE 73p.
PUB TYPE Reports - Research/Technical (143)

EDRS PRICE MF01/PC03 Plus Postage.
DESCRIPTORS Cognitive Processes; *Individual Differences; *Sex Differences; *Spatial Ability; Verbal Ability; *Visualization; Visual Measures
IDENTIFIERS *Spatial Tests; Structural Visualization

ABSTRACT

Two processing abilities used to solve spatial problems are examined: (1) the analog ability of structural visualization; and (2) the non-analog ability of verbal analytic reasoning. The distinction is based on an evaluation of information processing theory and a review of process-oriented studies of individual differences. Criteria are presented for determining which abilities are measured by tests of spatial ability and for classifying existing instruments on the basis of these criteria. Non-analog ability is the ability measured by certain tests of general intelligence and verbal processing abilities. Analog ability involves the holistic gestalt-like processing of visuospatial information. Reexamining the literature on individual differences and classifying the results by test types demonstrates that the commonly accepted belief that sex differences in spatial ability emerge in adolescence is not supported by the literature. The male advantage in analog processing ability is evident well before adolescence. Accumulated evidence also does not support the theory of recessive X-linkage of spatial ability. Five figures illustrate score differences, and 30 examples of spatial test items are appended. A 157-item list of references is included. (SLD)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

THE MEASUREMENT OF HUMAN VARIATION IN SPATIAL VISUALIZING ABILITY: A PROCESS-ORIENTED PERSPECTIVE

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☑ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

Robert F. Kyle

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

by
Michele F. Zimowski
and Werner Wothke

JOHNSON O'CONNOR RESEARCH FOUNDATION
HUMAN ENGINEERING LABORATORY

Technical Report 1986-1
September, 1986

TM012916

BEST COPY AVAILABLE

COPYRIGHT © 1986 BY JOHNSON O'CONNOR RESEARCH FOUNDATION, INCORPORATED

The Measurement of Human Variation in Spatial Visualizing Ability: A Process-oriented Perspective

Michele F. Zimowski and Werner Wothke

Abstract

In this report we distinguish two processing abilities used to solve spatial problems: the analog ability of structural visualization and the nonanalog ability of verbal analytic reasoning. Our distinction is based on an evaluation of information processing theory and results. It was motivated by the failure of factor analytic studies to adequately distinguish spatial tests on cognitive grounds. We present criteria for determining which abilities are measured by tests of spatial ability and classify existing instruments on their basis. The result is a clarification of certain inconsistencies in the individual differences literature. In view of our findings, we offer recommendations for an evaluation of the structural visualization measures currently in use at the testing centers of the Johnson O'Connor Research Foundation.

Table of Contents

Introduction 3

Psychometric Studies of Spatial Ability 3
- Early Validation Studies 4
- Factor Studies of the Spatial Domain 5

The Study of Visuospatial Information Processing 10
- Theories of Visuospatial Information Processing 11
- Studies of Spatial Item-feature Effects 12
- Task Attributes that Resist Nonanalog Processing 18

Classification of Tests of Spatial Ability 20
- Analog measures 20
- Nonanalog measures 22

Process-oriented Studies of Individual Differences 25

Conclusions 33

A Reexamination of the Individual Differences Literature 35
- The Developmental Trend of Sex Differences 35
- The Magnitude of the Sex Difference 37
- The Theory of Recessive X-linkage 38

Recommendations 40

Appendix 43

References 58

Introduction

Psychometric work has long supported a distinction between verbal and spatial skills but, as we will show, it has failed to provide an adequate explanation for the multiple factor solutions obtained in factor analytic studies of the "spatial domain." One must look to the information processing literature for clarification of the cognitive processes that differentiate spatial test performance.

In this literature we find evidence in favor of the conclusion that visuospatial information is processed in at least two distinct ways that depend in an orderly way on the item characteristics of the spatial task. Certain item features promote analog (holistic) processing of visuospatial information. Others promote nonanalog (verbal analytic) processing of visuospatial information. We use these item features as criteria to classify psychometric tests of spatial ability into two types: relatively pure measures of spatial ability (analog tests) and relatively impure measures of spatial ability (nonanalog tests). We find additional support for this classification scheme in process-oriented studies of individual differences. To demonstrate its utility, we examine results obtained in the individual differences literature by test type and find that the scheme clarifies certain inconsistencies. The results of this review have implications for the purification of existing measures of spatial ability and for the construction of tests that measure this ability in relatively pure form.

Psychometric Studies of Spatial Ability

Spatial ability has long been considered second only to verbal ability as an important aspect of human cognitive functioning. Galton (1883) was the first to establish its relevance before the turn of the century in his classical studies of the "faculty of visualizing." Tests of spatial aptitude were nevertheless omitted in the early factor analytic work of Spearman (1904) and Burt (1909), who claimed a uni-dimensional concept of "General Intelligence." Thorndike (1921) and McFarlane (1925) soon demonstrated that measures in the spatial domain were relatively independent of Spearman's

General Intelligence factor (G). Spatial measures were then routinely included in the multiple factor work of the 1920s and 30s (e.g., Kelley, 1928). In the studies reviewed by Wolfle (1940), the Spatial factor was second only to the Verbal factor in its frequency of occurrence.

Early Validation Studies

This growing evidence in favor of a distinct Spatial factor spurred interest in its practical significance: the use of spatial ability and its capacity to predict occupational success were extensively explored in the 1930s, '40s, and '50s in validation studies. Much of this work was conducted by the Johnson O'Connor Research Foundation with the Wiggly Block test. (See Daniel, 1978, for a detailed review of this work; see the Appendix for an illustration of an item from the Wiggly Block.)

This work revealed higher than average performance on the Wiggly Block in occupations that require visualization of solid form and average, or lower than average, performance in occupations traditionally associated with verbal skills. Groups of engineers (Technical Report 97; Validation Bulletin 74) engineering students (Technical Report 97), metallurgists (Validation Bulletin 11), and airplane mechanics (Validation Bulletin 22a) scored consistently above the male general population median on the test; groups of teachers and magazine editors scored consistently at or below this median (Statistical Bulletin 438; Statistical Bulletin 706; Technical Report 90). Scores on the Wiggly Block were also predictive of course grades in engineering (Brush, 1941), industrial arts (Remmers & Schell, 1933), and mechanical drawing (Technical Report 113), but they failed to predict rate of overall academic progress in high school (Technical Report 102) and high school grades in foreign languages, biology, chemistry, and English (Technical Report 113).

In comparative validation studies conducted outside of the Johnson O'Connor Research Foundation, measures of spatial and verbal abilities exhibited distinct patterns of correlations with technical proficiencies and academic success in various subject areas (see McGee, 1979, for a review).

Mechanical drawing ability, for example, correlated higher with spatial ability (r=.41-.66) than with verbal ability (r=.07-.26) in samples of engineer apprentices, shop students, and trade school apprentices (Holliday, 1940, 1943; Slater, 1941). Spatial scores were also better predictors of success in machine shop training and auto mechanics (Hunter, 1945; Martin, 1951).

In all, the results of the early validation studies support the distinction between verbal and spatial skills first observed in multiple factor studies. Spatial ability scores exhibit correlations with proficiencies not traditionally associated with verbal abilities, and higher than average spatial scores are found in occupations that require these proficiencies. Not surprisingly, these proficiencies typically involve the processing and retention of nonverbal information such as that associated with the arrangement of mechanical parts in an automobile engine or the plane drawing of three-dimensional forms.

Factor Studies of the "Spatial Domain"

The growing evidence in favor of a distinct Spatial factor also led to examinations of the factorial structure of the "spatial domain." Aided by the development of the centroid method of factor analysis and principal components analysis in the 1930s, the structure of the spatial domain was extensively explored in the 1940s and '50s. For these purposes, large batteries of "spatial" tests were independently constructed by different researchers. The numerous "spatial" tests nevertheless shared one characteristic: all required the processing of visuospatial stimuli.

To collate the findings of the three major studies that appeared in the literature (French, 1951; Guilford & Lacey, 1947; Thurstone, 1950), the Space and Visualization Committee was formed under the sponsorship of Educational Testing Service. A comparison of the results led the committee to propose a three-factor formulation (Michael, Guilford, Fruchter, & Zimmerman, 1957). *Kinesthetic Imagery*, a factor specific to tests requiring right and left hand discrimination, was specified as a minor factor; *Spatial Relations and Orientation* (SR-O) and *Visualization* (Vz) were designated

as primary factors and distinguished from one another in terms of the spatial reference point used in task solution. In the words of Michael et al. (1957):

> *The Spatial Relations and Orientation* factor represents an ability to comprehend the nature of the arrangement of elements within a visual stimulus pattern primarily with respect to the examinee's body as the frame of reference. In a typical test of the factor . . . the objects within the pattern hold essentially the same relationships to one another. In this rigid configuration there exists the implication that in perceiving the spatial arrangements of the stimulus pattern the respondent is able to distinguish whether one or more objects are higher or lower and/or farther or nearer than others, and/or whether a particular element may be to the right or left of another one (pp. 189- 190).
>
> *Tests of the Visualization* factor require mental manipulation of visual objects involving a specified sequence of movements. The objects appear within a more or less complex stimulus pattern. The individual finds it necessary mentally to rotate, turn, twist, or invert . . . objects, or parts, of a configuration . . . according to relatively explicit directions as to what the nature and order of manipulations should be. The examinee is required to recognize the new position, location, or changed appearance of objects that have been moved or modified . . . . Knowledge of whether the end results have been achieved satisfactorily is frequently apparent or implicit in that the examinee works through the manipulatory process until he completes the solution (p. 191).

In this account of the spatial domain, tests defining the *Spatial Relations and Orientation* factor include Aerial Orientation, Instrument Comprehension, Flags, Figures, Cards, Two-Hand Coordination, Lozenges, Cubes,

and the Guilford-Zimmerman Spatial Orientation test. Tests defining the *Visualization factor* include Surface Development, Form Board, Mechanical Movements, Copying, Punched Holes, Directional Plotting, and the Guilford-Zimmerman Spatial Visualization test. Sample items from most of these tests are shown in the Appendix.

Inspection of these items reveals that the cognitive distinction between primary factors is not entirely justified. Many tests thought to be representative of the SR-O factor appear to involve mental manipulation, an ability associated with the Vz factor. Thurstone's Lozenges A test, for example, loads on the SR-O factor even though instructions to mentally turn the figures are part of the test. Other measures of the SR-O factor also appear to involve mental rotation of two-dimensional figures. The Flags, Figures, and Cards tests of Thurstone and Thurstone (1941), for example, require the examinee to determine whether two drawings, usually presented in different positions, represent the same side of an object (see the Appendix). In all cases, the identity of the drawings can be determined by rotating one of the figures into congruence with the other. Tests defining each factor thus appear to involve the mental manipulation of figural forms with the observer removed from the stimulus pattern, an ability associated only with the Vz factor in the formulation of Michael et al. (1957).

The distinction among the major factors also proved tenuous on statistical grounds. Roff (1952) reported a correlation of .75 between the major factors which prompted Smith in his 1964 monograph (p. 92) to question the validity of the cognitive distinction among factors. Later studies based on confirmatory methods (Borich & Baumann, 1972; Price & Eliot, 1975) also failed to distinguish the primary factors, perhaps because of their small sample sizes.

The dimensionality of the spatial domain still remains an area of inquiry and discussion. Variants (e.g., Lohman, 1979; McGee, 1979) of the early two- and three-factor formulations continue to appear in the literature and there is little consensus about the number of factors and their cognitive meaning.

Based on his review of the individual differences literature, McGee (1979), proposes a two-factor formulation: Spatial Visualization and Spatial Orientation. According to McGee,

> *Spatial Visualization* is an ability to mentally manipulate, rotate, twist, or invert pictorially presented visual stimuli. The underlying ability seems to involve a process of recognition, retention, and recall of a configuration in which there is movement among the internal parts of the configuration, or of an object manipulated in three-dimensional space, or the folding or unfolding of flat patterns.
>
> *Spatial Orientation* involves the comprehension of the arrangement of elements within a visual stimulus pattern, the aptitude for remaining unconfused by changing orientations in which a configuration may be presented, and the ability to determine spatial relations in which the body orientation of the observer is an essential part of the problem. (pp. 3-4)

McGee's formulation is similar to the primary factors proposed by Michael et al. (1957) except that he explicitly differentiates two- and three-dimensional rotation tasks which define the Orientation and Visualization factors, respectively.

Lohman (1979) presents a different formulation which is based on his reanalysis of the spatial test data from a number of large studies conducted in the United States. He distinguishes three major factors: Spatial Relations, Spatial Orientation, and Spatial Visualization. In his formulation:

> *Spatial Relations* is defined by measures like the Cards, Flags, and Figures tests of Thurstone and Thurstone (1941). These tests require the encoding, matching, and rotation of relatively simple figures.
>
> *Spatial Orientation* is defined by measures such as Hands and the Guilford-Zimmerman Spatial Orientation test. These tests

are thought to measure an ability to imagine a stimulus array from different perspectives through attention to and translation of the orientation dimensions.

Spatial Visualization is defined by tests such as Paper-Folding, Paper Form Board, Surface Development, Hidden Figures, and Copying. These tests share two features: they are relatively unspeeded and they are more complex than tests defining the other two factors. Tests of this type often require the movement of parts of a stimulus configuration such as the folding or unfolding of a paper.

Sample items from most of these instruments are shown in the Appendix.

Lohman's account differs from other formulations in three ways. First, the Orientation factor is redefined and no longer includes two-dimensional rotation tasks but does cover the tasks that define the minor factor of Michael et al. (1957). Second, two- and three-dimensional rotation tasks are no longer distinguished; both define the Spatial Relations factor. Finally, tasks, such as the Paper-Folding test of French et al. (1963) and the Punched Holes task of Thurstone (1950), previously grouped with three-dimensional rotation tasks now define a separate factor, Spatial Visualization.

These inconsistencies among formulations can be explained, in part, by well-known limitations of the factor analytic approach. First, the methods used for factor extraction and rotation and the characteristics of the subject population and test battery affect the number of factors and their composition (see Carroll, 1983, for example). The studies of spatial ability have used different methods, subject populations, and test batteries. These differences undoubtedly contribute to the failure to reach consensus about the number of factors and their meaning. Second, factor studies based on overall test performance provide only limited information about the processes that underlie test performance. Factor analysts nonetheless attempt to describe these processes through inferences about features common to tests that load on the same factor. The factor descriptions that result do

not necessarily address the very features that influence task performance; tasks loading on the same factor share numerous properties, many of which are ignored in factor descriptions. Even when the relevant properties are identified, inferences about the underlying processes remain speculative. Factor studies do not directly clarify the nature of these processes and little is done to furnish the necessary experimental verification.

Perhaps because factor analytic attempts to distinguish spatial tests have not been entirely successful, the proposed classification schemes have been essentially ignored by other researchers. "Spatial" is still used rather indiscriminantly in the individual differences literature to refer to any test that requires the processing of visuospatial information. These "spatial" tests are rarely distinguished and commonly assumed to measure the same ability. The need for a classification scheme based on cognitive criteria is, however, evident. Results in the individual differences literature tend to be test-dependent. This is especially true of studies that have focused on identifying the biological and sociocultural determinants of the frequently observed sex difference favoring males on tests of spatial ability. Progress in this and other areas now depends on a better understanding of the abilities measured by the numerous "spatial" tests available for use by psychometrically oriented researchers. A finer level of analysis than that provided by factor studies of overall test performance is apparently required for these purposes.

The Study of Visuospatial Information Processing

A relatively new approach which developed independently of the individual differences literature holds promise for the clarification of the processes that underlie spatial task performance. In the last decade, spatial task performance at the *item level* has been extensively examined from an information processing perspective. The implications of this distinct approach for improved understanding, and thus measurement, of spatial ability have yet to be fully realized in the individual differences literature.

Theories of Visuospatial Information Processing

The motivation for this recent and extensive examination of spatial task performance at the item level is derived from the theoretical controversy surrounding the representation and processing of visuospatial information. The debate revolves around two related issues: (1) how visuospatial information is represented in memory, and (2) what type of processes operate on this memory. On the one hand, propositional theorists posit a monistic, amodal memory representation: visuospatial information in the form of percepts or images is encoded via abstract propositional structures, the same implicit structures used to encode verbal/linguistic information. On the other hand, spatial-imagery theorists, most notably Paivio (1969, 1977), postulate a dual code: the encoding format for visuospatial information is structurally and functionally distinct from that supporting verbal information. These dual coding theories typically assume that the code for verbal/linguistic information is propositional, or in some form suitable for verbal representation and processing, while the code for visuospatial information is analog in nature.

In accord with computer terminology, the term "analog" was initially used in the spatial-imagery framework to refer to the continuous nature of the processes and representations of visuospatial information (Moran, 1973). Because it proved difficult to distinguish between continuous change and a series of small discrete changes, this meaning has now been abandoned by most researchers (e.g., Baylor & Racine, 1977; Anderson, 1980). More recently, Paivio (1977) has used analog to indicate that the internal representation of visuospatial information bears some resemblance to the objects or things represented. Shepard and colleagues (e.g., Cooper, 1976a; Cooper & Shepard, 1973a, 1978; Metzler & Shepard, 1974; Shepard, 1978), however, use *analog* to denote a one-to-one correspondence between the intermediate states of the internal representation and those that the external referent would pass through while undergoing a similar transformation. Despite these differences in definition, in all cases *analog*:

1. implies "holistic" Gestalt-like processing,

2. refers to the assumed nature of the internal constraint that governs the form and use of visuospatial representations, and,
3. is used to differentiate the visuospatial system from that supporting verbal information processing.

Studies of Spatial Item-feature Effects

Evidence supporting analog processing models has been largely obtained by observing the effect of stimulus attributes on reaction time and error rates in transformation tasks. Although error rates are often recorded, reaction time (of correct responses) is the dependent measure of primary interest. Timed performance is examined as a function of stimulus properties to isolate the elementary operations involved in the solution of transformation problems and to clarify their nature.

The first empirical study of this type was carried out by Shepard and Metzler (1971). The transformation under study was mental rotation of the Shepard-Metzler block figures shown in the Appendix. Subjects were required to determine whether two block figures showed physically identical objects seen from different vantage points or in different directions.

When the same object was shown, Shepard and Metzler found that reaction time for a correct response was a linear function of the distance, in degrees, between the two orientations of the figure. This orientation effect on response time is illustrated in Figure 1. Cooper and Shepard (1973b) obtained similar results with alphanumeric characters as the stimuli. Similar orientation effects have since been reported in other groups of subjects and for other types of visuospatial stimuli (e.g., random shapes, Cooper, 1975, 1976a; embedded figures, Pylyshyn, 1979; PMA space figures, Mumaw, Pellegrino, Kail & Carter, 1984). These results are consistent with the analog interpretation of Shepard and colleagues.

More compelling support for this interpretation is found in process-monitoring studies where additional stimuli are presented during the transformation stage to monitor the mental rotation process (e.g., Cooper, 1976a;

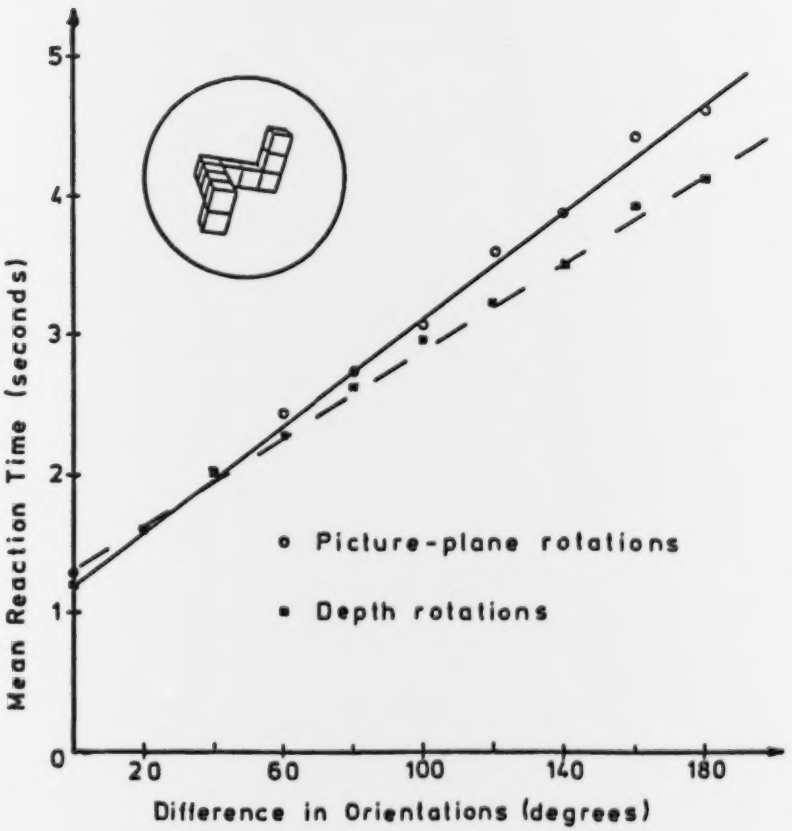


Figure 1: Mean response latencies for "same" judgments in the Shepard and Metzler (1971) mental rotation experiment. (From Shepard, 1984. Copyrights 1971 by the American Association for the Advancement of Science and 1984 by the American Psychological Association.)

Cooper & Shepard, 1973a,b; Metzler, 1973). Studies of this type are usually conducted in two phases. In the first phase, an empirical estimate of each subject's rotation rate is obtained by standard procedures. In the second phase, subjects are shown previously learned reference figures one at a time and instructed to mentally rotate them in designated directions. Test stimuli are then presented at predetermined times during the mental rotation stage in orientations that are compatible or incompatible with the presumed orientations of the representation. For each subject, the presumed orientation at the instant of presentation of a test stimuli is based on the previously obtained estimate of his or her rotation rate. The results obtained by Cooper (1976a) in a representative study of this type are presented in Figure 2. As the left hand panel shows, when the vantage point of the test stimuli is compatible with the presumed orientation of the representation ("probe-expected" trials), reaction times are relatively constant across orientation differences between the reference and test stimuli. As the right hand panel shows, when the orientation of the test stimuli differs from the presumed orientation of the representation ("probe-unexpected" trials), reaction times are a linear function of the distance, in degrees, between the orientation of the test stimuli and the presumed orientation of the representation. These results strongly suggest that during the rotation process a "holistic" representation of the stimulus passes through intermediate states that correspond to those a physical object would pass through while undergoing a similar transformation.

Other item-feature effects consistent with an analog model have been identified. Rotation rate is relatively independent of the plane of rotation (e.g., Metzler & Shepard, 1974; Shepard & Metzler, 1971) and unaffected when the representation of an object must pass through views in which parts of the object are not visible (Metzler & Shepard, 1974). Cooper and Podgorny (1976) also find that rotation rate, discriminative reaction time, and error rate are independent of the complexity of two-dimensional random figures, as would be expected under holistic analog assumptions.

Other workers have, however, identified item-feature effects that are inconsistent with an analog model. Pylyshyn (1979) and Yuille and Steiger

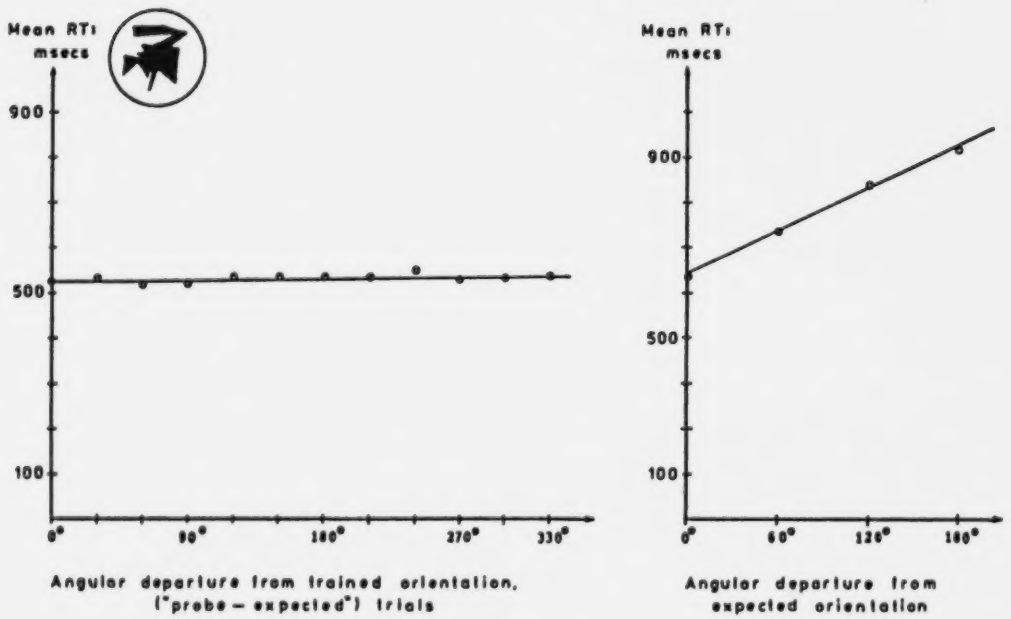


Figure 2: Response latencies from a "monitored" rotation experiment by Cooper (1976). The left hand panel shows mean reaction times as a function of the angular departure from the trained orientation when the probes are presented at expected orientations. The right hand panel shows reaction times when the probes are presented in unexpected orientations. (Copyright 1976 by the American Psychological Association.)

(1982), for example, report a complexity effect that depresses rotation rate by a factor between 2 and 6. (This result is in accord with certain propositional models that assume identity judgments are accomplished through piecemeal comparisons. In these accounts, the number of operations, and thus the overall reaction time required for solution, increases with the complexity of the stimulus.) The dimensionality of the stimuli also appears to affect rotation rate: Cooper (1975) and Cooper and Podgorny (1976) report rotation rates in excess of 230 degrees-per-second for two-dimensional random figures, while Shepard and Metzler (1971) report a 60-degree-per-second rate for three-dimensional figures. But as Shepard and Cooper (1982) have pointed out, procedural differences among the studies may account for these disparities; Podgorny (1975) and Cooper and Farrell (unpublished) fail to find dimensionality effects when rotation rates are obtained under identical administrative conditions. More compelling support for nonanalog processing is found in the work of Metzler and Shepard (1974). One of their eight subjects responded faster than expected under analog assumptions to stimuli that differed by 180 degrees. This subject reported using a nonrotational verbal strategy based on surface features to solve 180° rotation items.

Support for both types of processing is also found in the study of visuospatial transformations other than mental rotation. Studies of size-scaling tasks in which subjects are required to identify perceived and remembered objects differing in physical size are one example. Many workers using different types of visuospatial stimuli have found a linear relationship between reaction times for judgments of shape identity and the physical differences of the two stimuli expressed in ratio form (e.g., Bundesen, Larsen, & Farrell, 1982; Bundesen & Larsen, 1975; Larsen & Bundesen, 1978; Sekuler & Nash, 1972). Their results are in accord with an analog model. Nevertheless, the failure of others (e.g., Kubovy & Podgorny, 1981) to replicate the linear effect suggests that size-scaling problems may also be solved through nonanalog means.

Neither analog nor propositional models with predictions based on the *basic nature* of their respective codes can accommodate this evidence in fa-

vor of both types of processing. Propositionalists have, however, modified their theories to incorporate analog effects; analog predictions are based on arbitrary constraints placed on the system rather than on the nature of the system's propositional code. This virtually unlimited capacity of the propositional approach to account for analog effects through arbitrary constraints has led some researchers, most notably Anderson (1976, 1978, 1979), to conclude that the issues of representation cannot be resolved on strictly empirical grounds.

Despite contradictory evidence, some researchers (e.g., Paivio, 1977) continue to advocate structurally and functionally distinct forms of representations for *all types* of verbal and visuospatial information. Others (e.g., Shepard, 1981) concede that the analog processes associated with visuospatial information processing are not intrinsically incompatible with a propositional code provided different organizational principles distinguish this type of processing from linguistic and other nonanalog information processing in the system. Still others propose integrative models incorporating features from both propositional and spatial-imagery theories (e.g., Baylor & Racine, 1977; Funt, 1983; Kosslyn, 1981; Kosslyn & Shwartz, 1977). Kosslyn and Shwartz (1977), for example, propose that two types of representations underlie visuospatial information, one propositional in nature and representing conceptual information, the other perceptual in nature and representing appearances.

Even though examinations of spatial item-feature effects have not resolved the theoretical controversy, the results of these studies have implications for the measurement and definition of spatial ability. Certain task features are consistently associated with analog effects in these studies, while others are associated with nonanalog effects. Still others are associated with both analog and nonanalog effects. This pattern suggests that: (1) visuospatial problems are solved in at least two different ways–through analog and nonanalog processes, and (2) the type of solution process is predictable on the basis of task selection. The implications are that: (1) *psychometric tests* of spatial ability may measure either or both processing abilities, depending upon the characteristics of their items, and (2) the

identification of task characteristics that resist nonanalog processing provides a means for differentiating spatial tests.

Task Attributes that Resist Nonanalog Processing

Tasks shown to resist nonanalog processing in studies of item-feature effects share a number of properties:

1. *The tasks involve judgments among rotated stimuli.* Other transformation tasks such as size-scaling are less resistant to solution by nonanalog processes (e.g., Kubovy & Podgorny, 1981).
2. *The stimuli differ by orientations other than 180 degrees.* Items with 180-degree rotations may be solved through nonanalog strategies even by subjects who readily solve rotation items through analog methods. As previously discussed, Metzler and Shepard (1974) report shorter response times than expected under analog assumptions for 180-degree items for one of their eight subjects. This subject reported using nonrotational strategies. Apparently because simple verbal rules such as "the left side will become the right side" are readily applied, 180-degree rotation items are less resistant than others to nonanalog processing. These results have implications for any "spatial" task in which simple verbal rules can be substituted for the rotation or transformation process. These tasks will exhibit item-feature effects favoring a nonanalog model of visuospatial information processing.
3. *The distractors of the rotation tasks are mirror images of the reference stimuli or structurally equivalent forms.* When the two stimuli of an item are physically different they differ only by a reflection. This procedure minimizes the effectiveness of nonanalog "feature-extraction" strategies in the solution of mental rotation problems. When

mirror-image distractors are not used, the problems are readily solved through identification of incongruent portions of the figures (see Zimowski, 1985). In this connection, Cooper (1976b) has shown an effect of similarity of test and reference forms on discriminative response latencies. For certain subjects classified as Type II, "different" latencies are a monotonic function of stimulus similarity. Based on this pattern, Cooper (1976b) has suggested that Type II subjects analytically compare the two stimuli and check for a mismatch through location of distinctive features. The results on which this conclusion is based are well-replicated (Cooper & Podgorny, 1976; Cunningham, Cooper, & Reaves, 1982; also see Cooper 1980a,b).

4. *The items require whole-whole rather than part-whole or part-part comparisons.* Tasks such as those used by Pylyshyn (1979) and Palmer (1977) (see the Palmer-Pylyshyn figures in the Appendix) that require subjects to determine whether rotated probes are true subfigures of reference stimuli produce effects consistent with a nonanalog model of visuospatial information processing (see Pylyshyn, 1979, for example). In accord with the reaction time results, subjects report using serial comparison and other nonanalog strategies on items of this type.

   These results have implications for tasks that ostensibly require whole-whole comparisons but contain figures that may be readily parsed into distinguishable, nonredundant subfigures. Tasks of this type will also be more susceptible to nonanalog processing because they can be solved through serial comparisons of subfigures. Although these serial comparisons may involve analog manipulation of subfigures, other nonanalog processes, such as figure parsing, and subfigure disembedding will be involved.

   Part-by-part comparison strategies can be inhibited through use of mirror-image distractors and stimuli that

yield subfigures of similar structure when they are parsed. Under these conditions, a parsing strategy will presumably lead to errors in the comparison process.

5. *The items require the rotation of an entire object as a rigid whole rather than the rotation of only one or several pieces of the object relative to the whole.* The work of Shepard and Feng (1972) with items resembling the paper-folding and surface development tasks of spatial batteries (e.g., French, Ekstrom, & Price, 1963) suggests that items of this type may be solved through analog processes by some subjects. At the same time, the degree of intersubject variability suggests that other subjects use nonanalog solution processes.

Classification of Tests of Spatial Ability

On the basis of the item features outline above, tests of spatial ability were classified into two types: analog and nonanalog. Analog tests contain items with features that resist nonanalog processing and are assumed to measure holistic processing in relatively pure form. Nonanalog tests contain items with features that are susceptible to nonanalog processing. The items of these tests are assumed to be solved in part, or in whole, by some subjects through logical reasoning or other faculties of the verbal domain.

The result of the classification is shown below. The lists are not exhaustive but contain many of the tests currently in use by psychometrically oriented researchers. Sample items from most of these tests are presented in the Appendix.

Analog Measures

1. *Vandenberg-Shepard Mental Rotations Test* (Vandenberg, 1971). This test is a multiple choice version of the Shepard-Metzler figures constructed for group administration. The

subject is required to select the two of four response alternatives that are rotated versions of the reference stimuli. Most of the distractors are mirror images of the reference stimuli, and a time limit is placed on each of the two 10-item subsets of the test. (Six of the items in this test may be efficiently solved through nonanalog means; both of the distractors in each of these items are nonmirror images of the correct alternative. Nevertheless, the test is assumed to be a relatively pure measure of spatial ability because the remaining 14 items have properties that resist solution through nonanalog processing.)

2. *The Guilford-Zimmerman Spatial Visualization Test* (Guilford & Zimmerman, 1947). In this test, subjects are asked to mentally rotate a picture of a clock in a specified direction and select the alternative that shows the clock in its final position. Each alternative is a picture of the clock as viewed from a different perspective. As a result, the alternatives do not contain distinctive features that allow for rapid elimination of the incorrect alternatives. In the Bock-Kolakowski (1973) modification of this test, the items are presented on slides and their exposure times are individually controlled.

3. *The Analog Subset of the Incomplete Open Cubes Test* (Zimowski, 1985). In this test, subjects are asked to determine whether two incomplete open cubes fit together to form a complete open cube. A rotation of one or the other cube is apparently required to determine the compatibility of the cubes in the subset of items especially constructed to inhibit analog processing. This subset contains stimuli that are not readily parsed into nonredundant subfigures for piece-by-piece comparisons. The items of this subset also differ by orientations of other than 180 degrees and the distractors are mirror images of the compatible configurations. The items are presented on slides and their exposure times are individually controlled.

The items of each of these analog measures have properties that resist nonanalog processing. Thus, they are not efficiently solved through verbal reasoning strategies. Nevertheless, strategies of this type can undoubtedly be used to solve these problems if enough time is allowed for their application. For this reason, time restrictions have been imposed on each of these tests to inhibit solution through other than analog means.

This "speeded" characteristic of psychometric measures of analog ability apparently contributes to their purity. Its effect on spatial task performance has not been explicitly examined in the information processing literature, nor has the absence of time restrictions encouraged solution through nonanalog means. The apparent contradiction can be explained by paradigmatic differences between the two approaches. In most of the information processing studies, variation in analog ability (which is the focus of the psychometric approach) has been deliberately minimized through extensive training or selection of subjects for homogeneous aptitude. In particular, subjects adept at analog processing are selected. Spatial tasks are thus solved through analog means by these subjects unless task properties promote nonanalog strategies by permitting efficient solution through their use. In examinations of individual differences, on the other hand, samples of heterogeneous ability are selected because variation in analog processing ability is of interest. Subjects with low ability in these samples are likely to resort to nonanalog strategies, even when the tasks can not be efficiently solved in this manner. Speeded conditions are thus required to inhibit nonanalog solution strategies even when the item properties resist their application.

Nonanalog Measures

The list below includes examples of nonanalog measures. Some of these instruments were explicitly constructed to measure nonanalog processing of visuospatial information (i.e., the nonanalog subset of the Incomplete Open Cubes test, Raven's Progressive Matrices). Others were apparently constructed to measure analog processing in relatively pure form but fail to do so because of their item properties (e.g., DAT Space Relations Subset).

Still others were intended to assess other aspects of cognitive performance (i.e., field independence-field dependence, Embedded Figures test). Examples of each are included to illustrate the range of nonanalog tests often referred to as "spatial" in the individual differences literature. Most of the nonanalog tests described below are included in the collection of "spatial" measures recently compiled by Eliot and Smith (1983).

1. *The Nonanalog Subset of the Incomplete Open Cubes Test* (Zimowski, 1985). This subset of items was especially constructed to encourage nonanalog processing of the stimuli. The items contain distinctive features that permit solution without rotation. Nonmirror-image distractors are used.

2. *The Space Relations Subtest of the Differential Aptitude Test Battery* (Bennett, Seashore, & Wesman, 1974). In this test, subjects are presented with an unfolded version of an object and asked to select the one of four alternatives that is the folded form. The objects include rectangles, squares, octahedrons, cube-based pyramids, and other three-dimensional geometric shapes. The distractors of each item differ from the correct alternative in structure or in the pattern of shading of the component parts.

   Items in this test have a number of properties that promote solution through nonanalog strategies. In many cases, the detection of structural nonequivalence does not require the mental folding of the reference figure. Nonequivalence can be detected simply by noting a feature (or side) of the unfolded version that is absent in the distractors. Structurally equivalent distractors can often be eliminated in a similar manner by noting a surface that is shaded (or unshaded) in the reference figure but unshaded (or shaded) in the folded forms.

3. *Embedded Figures* (Witkin, 1950). The Embedded Figures test requires subjects to locate simple geometric shapes within complex visuospatial configurations. The task thus

requires part-whole comparisons and disembedding, both of which were shown to involve nonanalog processing (e.g., Pylyshyn, 1979).

4. *Raven's Progressive Matrices Test* (Raven, 1938). The items of this test contain arrangements of visuospatial stimuli that differ in quantitative and qualitative dimensions. In each item, a segment of the arrangement is missing. Subjects are asked to select the alternative that completes the arrangement. The items do not require analog processing for their solution. They do not contain any of the properties shown to inhibit nonanalog processing nor any that promote or require analog processing. Instead, some of the items require perceptual accuracy. Others require an understanding of the logic of the spatial structure (see Raven, 1938).

   The test was designed to measure Spearman's G factor in a culture-free manner (see Anastasi, 1958) rather than spatial ability. Even so, it is often considered to be a measure of spatial aptitude (see Caplan, MacPherson, & Tobin, 1985).

5. *Minnesota Paper Form Board Test* (Likert, 1934; Likert & Quasha, 1970). In the items of this test, the reference stimulus is a set of five disjoined pieces. The subject is asked to select the figure (one of five alternatives) that can be constructed from the set of pieces. The constituent parts of the completed figures are in different orientations than the respective pieces of the reference stimulus. The task is thus assumed to require mental rotation for its solution.

   Items in the test, however, have a number of properties that encourage nonanalog strategies. First, the items require part-whole rather than whole-whole comparisons. Second, some of the distractors can be eliminated simply by noting a piece that is present in the reference stimulus but absent in the figure. In many cases, this feature-

extraction strategy is very effective because the pieces of the distractors and the reference stimulus are very dissimilar in structure. In other cases, the pieces of the distractors and the reference stimulus differ only in size. These distractors can be eliminated by simply detecting differences in stimulus size. Size discrimination is not necessarily an analog process, however.

6. *Paper-Folding Test* (French, Ekstrom, & Price, 1963). In each item of the paper-folding test, the subject is shown a series of figures obtained from folding a square sheet of paper and punching a hole in the folded form. Each step of this folding and punching process is depicted in a separate figure of the series. The subject is asked to determine the position of the holes if the paper were unfolded. The subject either selects the one of five alternatives that shows the correct position of the holes in the unfolded form, or, as in the version in use at Johnson O'Connor, indicates the position of the holes on a square grid or sheet of paper.

   For many of the items in versions of the paper-folding test, simple verbal rules can be substituted for the mental-folding process. The item shown in the Appendix (from Thurstone's Punched Holes test), for example, can be readily solved through application of a symmetry principle.

Process-oriented Studies of Individual Differences

Additional support for this classification of spatial tests is found in process-oriented studies of spatial test performance where individual differences are analyzed through application of cognitive process theory and methodology. This relatively new approach to the study of cognitive abilities has been referred to elsewhere as componential analysis (Sternberg, 1977) and cognitive components research (Pellegrino & Glaser, 1979). Its goal is to overcome the limitations of factor analysis and other psychometric methods by analysing cognitive test performance in much the same way

that cognitive psychologists examine cognitive tasks–through use of process models.

In cognitive components research, the process models used to describe the components involved in the solution of cognitive test items are often based on those developed in the information processing literature. The difference is that individual variation in the components of the model and its relationship to performance on psychometric tests are of interest in cognitive components research. To assess this variation, samples of heterogeneous aptitude are selected, laboratory tasks are modelled after test problems found in psychometric reference tests, and performance on the reference test is often recorded. The laboratory tasks typically vary in systematic ways that allow for testing of the process model and for separation of the variation attributable to the model components. This variation is often correlated with performance on the reference test to assess the contribution of each component to psychometric test performance. Research of this type is thus an integration of the psychometric and information processing approaches to the study of cognitive abilities. As we shall show, it provides a link between these literatures and supports the proposed classification of spatial tasks that was developed entirely on the basis of information processing results.

The bulk of this type of research in the area of spatial ability has been performed by Pellegrino and colleagues (e.g., Mumaw & Pellegrino, 1984; Pellegrino, Alderton, & Regian, 1984; Pellegrino, Mumaw, & Shute, 1985) and Snow and colleagues (e.g., Kyllonen, Lohman, & Snow, 1984; Lohman & Kyllonen, 1983; Snow, 1978). Most of this work has been influenced by the distinction among "spatial" factors forwarded by Lohman (1979) and his conclusion that these factors represent two correlated indices: speed-power and simplicity-complexity. According to Lohman (1979), the *speed-power* index reflects the amount of time required for solution and the relative emphasis placed on speed and accuracy of responding. (A rough index of "speededness" is the number of test items divided by the total amount of time allowed for the test.) The *simplicity-complexity* index, on the other hand, reflects the complexity of the task stimuli and that of the cognitive

processes required for solution. (A crude measure of complexity is the dimensionality of the stimuli and/or the number of separate elements that must be processed for solution.) Together these two indices are assumed to reflect the differences among spatial measures that contribute to their differentiation in factor analyses.

The differences among the three types of spatial tests distinguished by Lohman (1979) are reflected in their relative positions on the two indices. The ordering on the speed-power index indicates that Spatial Relations problems are solved most rapidly. Spatial Orientation problems closely follow but Spatial Visualization problems require considerably more time for their solution. The ordering on this index also reflects the relative emphasis placed on speed and accuracy by the administrative formats of the tests. The ordering on the simplicity-complexity index, on the other hand, indicates that Spatial Visualization problems contain the most complex stimuli and require the largest number of mental operations for their solution. Although there is complexity variation among Spatial Relations problems (2- vs. 3- dimensional problems), these tasks, as a whole, are less complex than Visualization problems. The stimuli have fewer elements and the problems typically require only one spatial transformation. Spatial Orientation problems, as a whole, are also less complex than Visualization problems. Their position on the simplicity-complexity continuum is the same as that of Spatial Relations problems.

If Lohman's formulation is correct, the individual variation measured by Spatial Relations tests with simple two-dimensional stimuli should relate to process estimates of speed rather than accuracy. The variation measured by Spatial Visualization tests and more complex Spatial Relations problems, on the other hand, should relate to both. The more complex tasks should also require a larger number of component processes and/or multiple executions of the same process.

Support for this view is found in process-oriented studies of Spatial Relations where the process model of Shepard and colleagues (see Shepard & Cooper, 1982) has been applied. This model assumes that the solution pro-

cess consists of at least four sequential components: the encoding, rotation, comparison, and response stages. Processing begins in the encoding stage where representations of the stimuli are encoded. In the rotation stage which follows, one of the two representations is rotated into congruence with the other. Next, in the comparison stage, the stimuli are examined to determine if they represent the same configuration. Finally, in the response stage a same or different response is executed. Estimates of the processing parameters of this model are computed for each individual by regressing the item response latencies of that individual on the number of degrees of rotation required to solve each item. The slope of this function is used as an index of rotation rate for that individual, the intercept as an index of the encoding, comparison, and response processes.

Mumaw, Pellegrino, Kail, and Carter (1984) applied this model to two types of stimuli: familiar alphanumerics and unfamiliar shapes. The unfamiliar shapes were similar to those found in the PMA Spatial Relations instrument which served as the psychometric reference test in the study. Four process measures, a slope and intercept for each type of task, were computed for each subject.

A number of interesting findings emerged. There were only small differences in the intercepts of the response functions of the familiar stimuli. For the unfamiliar shapes, on the other hand, subjects varied widely in the time required to encode, compare, and respond. These differences were related to scores on the reference test. The slope estimates obtained from both types of stimuli were also related to scores on the reference test; less steep slopes were associated with higher ability. Error rates for both types of tasks, on the other hand, were low and unrelated to spatial test performance.

In a related study, Pellegrino and Mumaw (1980) used a three-dimensional rotation task and were able to replicate the relationship between rotation rates, intercepts, and spatial ability as measured by a reference test. Error rates were also related to spatial test performance. Similar results have been reported by Egan (1978, 1979) and Lansman (1981). Egan

(1978), however, failed to replicate the relationship between rotation rate and spatial test performance.

Pellegrino, Alderton, and Shute (1984) interpret these findings as support for Lohman's (1979) factorial distinction. Psychometric Spatial Relations tests with simple two-dimensional stimuli measure individual differences in speed of encoding, rotating, and comparing stimuli rather than the accuracy of these processes. More complex Spatial Relations tests with three-dimensional stimuli measure individual differences in both speed and accuracy.

Another interpretation of these results is, however, just as compelling. The laboratory tasks with two-dimensional stimuli were simply not difficult enough to discriminate among subjects in right-wrong performance when the tasks were presented under unspeeded conditions. Under this administrative format, most subjects were able to solve the items correctly as indicated by the low error rates obtained by Mumaw et al. (1984). Because of this floor effect there was little variation in error rates and therefore they failed to correlate with spatial test performance. If more difficult items with two-dimensional stimuli are included in laboratory tasks, error rates may also correlate with spatial test performance.

Process-oriented studies of Spatial Visualization also appear to support Lohman's interpretation. These studies have thus far examined performance on laboratory tasks modeled after the Paper Form Board and Paper-Folding items of psychometric tests. In a representative study of this type, Mumaw and Pellegrino (1984) examined performance on a laboratory task modeled after items from the Minnesota Paper Form Board test. The task was designed to test a process model with five sequential components: the encoding, search, rotation, comparison, and response components. As shown in Figure 3, some of the items were constructed to require only a subset of the components for their solution. Others were constructed to include all of the components of the model. This item manipulation was introduced to obtain estimates of the time required for each component.

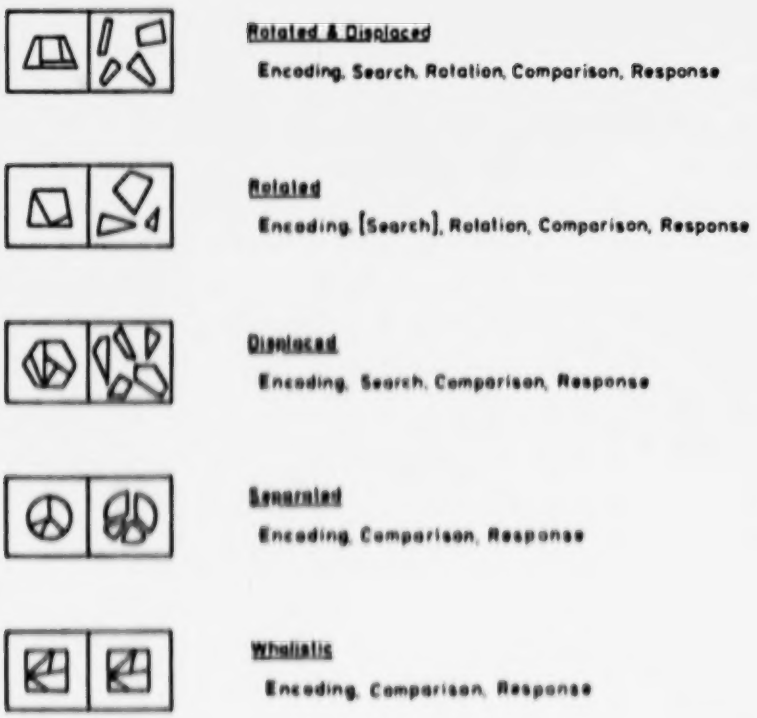


Figure 3: Component analysis for Paper Form Board tasks (from Pellegrino, Alderton, and Regian, 1984).

The other item manipulation, the number of stimulus elements, was included to test the hypothesis that the relevant processes were repeatedly executed for each stimulus element.

The results of this study are consistent with Lohman's interpretation. Performance on positive trials, where the elements matched those of the completed figure, was a function of stimulus and processing complexity. Within each problem type shown in Figure 3, solution time and error rates increased with the number of stimulus elements. Both performance measures also increased with the number of distinct processes required to complete the problem. Also in accord with Lohman's account, both speed and accuracy were predictive of performance on the reference test, with accuracy making the largest contribution. Similar findings have been reported by Mumaw et al. (1980) and Pellegrino et al. (1981).

Although these findings support Lohman's formulation, we conclude that tests of Spatial Visualization load on a separate factor not simply because they are relatively unspeeded and require more complex processing due to their stimulus attributes, but because these characteristics promote or require solution through nonanalog processes. In support of this conclusion, we have shown that tests of this type (e.g., the Paper-Folding test of French et al., 1963, and the Minnesota Paper Form Board test) contain items with features associated with nonanalog effects in the information processing literature.

Additional support for our view is found in process-oriented studies of solution strategies. In an analysis of the form board task, Pellegrino, Mumaw, & Shute (1985) report systematic differences between the response functions of subjects who perform poorly or average on the reference test and subjects who exhibit superior performance. These differences suggest that subjects of low and intermediate ability employ wholistic strategies and merge the pieces of the stimulus set before comparison with the completed figure, while subjects of high ability use a more analytic (piece-by-piece) comparison strategy. (This piece-by-piece strategy may or may not involve an analog component in the rotation stage; the relationship between the

angular disparity of the respective pieces of the completed figure and the stimulus set was not examined). The use of analog strategies is thus associated with poorer performance on this so-called spatial test.

Strategy differences among individuals have also been reported on other measures of Spatial Visualization. Snow (1978, 1980) and Kyllonen, Lohman, and Snow (1984) report both feature-extraction and mental-construction strategies in the solution of items from the Paper Folding test of French et al. (1963). In a paper-folding task modeled after the Differential Aptitude Space Relations subtest, Alderton and Pellegrino (1984) also note individual differences in the response functions indicative of distinct strategies. Tests of Spatial Visualization thus measure abilities other than analog processing.

Information about the nature of these abilities is provided by correlational studies of Spatial Visualization. Kyllonen, Lohman, and Snow (1984) report a correlation (r=.42) between performance on a Visualization test, paper-folding, and verbal ability, while Daniel (Technical Report 1983-2) notes correlations in the range of .39 to .47 between performance on the Paper-Folding test and the Analytical Reasoning test. Similary, Bennett et al. (1974) report substantial correlations among the Space Relations subtest of the Differential Aptitude Test battery and the language usage (r=.54) and verbal reasoning (r=.60-.67) subtests of the same battery. Moreover, Snow et al. (in press) were unable to separate Spatial Visualization from general Fluid ability (Horn & Cattell, 1966) in a factor analysis. The nonanalog processes used to solve Spatial Visualization items thus appear to be the same linguistic abilities assessed by certain tests of general intelligence and verbal abilities.

We find additional support for this conclusion in the distributional properties of analog and nonanalog abilities. Scores on nonanalog tests exhibit distributional properties which are distinct from those associated with analog ability but similar to those associated with verbal processing abilities. In an extensive meta-analysis of existing studies of spatial aptitude that was based on Lohman's distinction among factors, Linn and Petersen (1985)

find that males and females perform about equally well on tests of Spatial Visualization. On the measures of mental rotation included in their study, which we would classify as analog tests, they report a consistent sex difference favoring males. Using item-factor analysis (Bock et al., 1985) and facet design methodology (Thissen, 1982), Zimowski (1985) reaches much the same conclusion in her process-oriented examination of spatial test performance. Large and consistent sex differences favoring males are associated with relatively pure measures of analog processing (i.e., Analog subset of the Incomplete Open Cubes test; the Vandenberg-Shepard Mental Rotations test). Those associated with measures promoting or requiring nonanalog processing are of much smaller magnitude (i.e., Nonanalog subset of the Incomplete Open Cubes test), nonexistent (Differential Aptitude Test-Space Relations subtest, Factor 2 of Raven's Progressive Matrices), or favor females (Factor 1 of Raven's Progressive Matrices).

Zimowski (1985) also notes consistent differences in the distributional form of the abilities associated with each test type. Only the score distributions of the analog measures exhibit bimodality in a consistent manner. Those of nonanalog measures are, for the most part, normally distributed. Thus, only analog processing ability is in accord with a major gene hypothesis that has been proposed to explain the qualitative distinction between good and poor visualizers (see Bock & Kolakowski, 1973, or O'Connor, 1943). Nonanalog processing ability, on the other hand, is in accord with a polygenic model of inheritance also associated with verbal abilities.

Conclusions

From our review of the information processing literature and process-oriented studies of individual differences, we conclude that nonanalog and analog visuospatial processing abilities represent distinct sources of individual variation. Nonanalog ability is the same ability measured by certain tests of general intelligence and verbal processing abilities. Raven (1938) was perhaps the first to realize this fact when he constructed a test consisting entirely of visuospatial information to measure Spearman's G in a culture-free manner. Analog ability, on the other hand, is distinct from

verbal abilities and involves holistic gestalt-like processing of visuospatial information. It is this ability that displays a sex difference favoring males and bimodality in the within-sex distributions.

Our distinction between analog and nonanalog visuospatial processing abilities is not entirely new to the individual differences literature. As early as 1950, Spearman and Jones noted that:

> in the sense that attention wanders from one element of the figures to another. The other mode of operation is comparatively synthetic, in that the figures (or their constituents) are mentally grasped in much larger units (sometimes called "wholes"). The former procedure, not the latter, tends to load noegenetic [i.e., congeneric] processes with G (p.70).

More recently Maccoby and Jacklin (1974) have distinguished two types of spatial ability: nonanalytic and analytic. The former, measured by mental rotation tests such as the Guilford-Zimmerman and the Vandenberg-Shepard, is assumed independent of verbal ability; the latter, measured by instruments such as the Embedded Figures test, is influenced by verbal skills.

What differentiates our analog-nonanalog distinction from similiar distinctions in the literature is that it is based on item-feature effects found in information processing studies. We show that the extent to which tests with visuospatial content measure analog and nonanalog processing abilities depends in an orderly way on the item characteristics of the task. Only tests with properties that resist nonanalog processing measure analog processing ability in relatively pure form. Other tests without these properties measure nonanalog processing in part or in whole.

All tests of the latter type have been classified as nonanalog tests in our scheme. Performance on some of these tests will undoubtedly be aided by spatial (analog) cognition if their items can also be solved through analog means. Performance will, however, tend be a mixture of verbal and spatial strategies and thus confound the results of studies of individual differences.

Because only analog tests measure an ability that is relatively distinct from the abilities assessed by tests with explicit verbal content, we conclude that only analog measures should be called *tests of spatial ability*.

The indiscriminate classification and use of impure (nonanalog) measures of spatial ability have led to inconsistencies in the individual differences literature. The use of impure measures has also led to conclusions that are challenged by closer inspection of the item properties of the tests.

To illustrate our point, in the next section we classify results obtained in the individual differences literature by test type. In our first application of the classification scheme, we show that the commonly accepted belief that sex differences in spatial ability emerge at adolescence is not supported by the literature when the content of the tests typically used to assess preadolescent sex differences is examined according to our cognitive criteria. In our second application, we find that when only relatively pure measures of spatial aptitude are considered, the sex effect favoring males is not small or inconsistent, as Hyde (1981) and Caplan, MacPherson, and Tobin (1985) have recently concluded. The final application reveals that even when only relatively pure measures of spatial ability are considered, the accumulated evidence does not support the theory of recessive X-linkage. This conclusion is in accord with that reached in the literature.

A Reexamination of the Individual Differences Literature

Application 1: The Developmental Trend of Sex Differences

It is generally believed that sex differences in spatial ability do not emerge until adolescence. This view was made popular by Maccoby and Jacklin's 1974 review of published studies. A sex difference favoring males on tests of nonanalytic spatial ability was found in only one of 22 studies with preadolescent samples but a male advantage was reported in eight of the nine studies with adolescent or adult samples.

Since this influential review, theoretical efforts have considered this developmental aspect a fact that must be explained. Waber (1977), for example, suggests that the timing of maturation through its influence on brain lateralization for spatial functions accounts for the sex difference favoring males that emerges during adolescence. In her theory, females, in comparison with males, are assumed to be early maturers since both the initiation of hormonal events preceding puberty and peak height velocity occur earlier. Because of this, females have less developmental time than males to become lateralized for spatial functions. As a result, they exhibit poorer spatial skills from maturity onward. Similarly, other theorists have attributed the sex-related difference to the increase in androgen levels that occurs in males at adolescence (e.g., Bock, 1973; Petersen, 1976).

An examination of the tests typically used with preadolescent samples indicates that these theoretical efforts have been misguided. Most of the tests used in the studies reviewed by Maccoby and Jacklin (1974) and in more recent work reviewed elsewhere (see Petersen et al., 1982) fall into Lohman's Spatial Visualization category and have nonanalog properties. Several of these studies used the DAT Space Relations subtest (Connor, Schackman, & Serbin, 1978; Nash, 1975; Sherman, 1980), a nonanalog test that is associated with negative or inconsistent results in adult samples (Tapley & Bryden, 1977; Zimowski, 1985). Others used Wechsler's Block Design test (Regard, Strauss, & Knapp, 1982; Strauch, 1976; Taylor, 1977), another nonanalog instrument that also shows negative or inconsistent results in adult samples (e.g., Petersen, 1976).

In the few studies using tests with properties known to reveal sex differences in adult samples, a difference favoring preadolescent males has been observed (Levine & Huttenlocher, 1985; Richmond, 1980; Rosser, Ensing, Glider & Lane, 1984). Rosser et al., for example, report the superior performance of 4- and 5-year old boys on a mental rotation task with properties known to inhibit nonanalog processing. Similarly, Richmond (1980) notes a sex difference favoring 9½- and 10½-year-old boys on the PMA Spatial Relation test.

The same pattern of sex differences by test type first observed in adult samples (see Zimowski, 1985) is thus apparent in preadolescent samples as well. The implication is that the biological changes associated with the onset of adolescence are not responsible for the male advantage in analog processing ability. This advantage is evident well before the physical changes that accompany adolescence are initiated.

Linn and Petersen (1985) reach much the same conclusion in their meta-analysis of studies conducted since Maccoby and Jacklin's (1974) review and before June of 1982. They observe no change in the size of the male advantage at early adolescence as measured by the PMA Spatial Relations test.

Application 2: The Magnitude of the Sex Difference

Based on reviews of the literature, several workers (e.g., Caplan, MacPherson, & Tobin, 1985; Hyde, 1981; Plant, Southern, & Jacklin, 1977) have concluded that the sex difference favoring males on tests of spatial ability is small or nonexistent. Hyde (1981), for example, reanalyzed the studies reported by Maccoby and Jacklin (1974) and found that only one to five percent of the variance in spatial scores could be attributed to sex. Moreover, the differences between overall male and female performance were only about one-fourth to one-half of a standard deviation.

Tests that measure nonanalog processing were, however, indiscriminantly included in this and in other reviews. The sex differences exhibited by impure tests of this type should be small, nonexistent, or favor females because these tests measure verbal analytic abilities. Only tests that measure analog spatial ability in relatively pure form should reveal substantial male advantages. The meta-analysis of Linn and Petersen (1985) illustrates this point. They report an effect size of about 1 standard deviation for the Vandenberg-Shepard Mental Rotations test, a relatively pure measure of spatial (analog) ability. On impure measures, which are classified as tests of Spatial Visualization, males and females perform about equally well. Zimowski (1985) reports similar findings by test type.

The implication is that any review that includes a representative sample of published studies and does not classify the results by test type will conclude that the male advantage is small. This is because most of the studies in the literature have used impure measures of spatial ability. Of the studies conducted between 1974 and June of 1982 (see Linn & Petersen, 1985), for example, nine used relatively pure measures of mental rotation, while 19 used relatively impure measures.

Application 3: The Theory of Recessive X-Linkage

The theory of X-linkage was first proposed by Johnson O'Connor (1943) to explain the male advantage on spatial tests and the bimodality commonly observed in the within-sex spatial score distributions. The theory predicts (with the assumptions of random mating and a gene frequency of .5) that one-half of males but only one-fourth of females will possess the trait. (See McClearn & DeFries, 1973, for a discussion of X-linkage). Many studies have yielded proportions in the upper and lower components of the score distributions consistent with these predictions (e.g., Bock & Kolakowski, 1973; Zimowski, 1985). The proportions (p's) obtained by Zimowski (1985) for the Vandenberg-Shepard Mental Rotations test are shown in Figure 4.

The theory of X-linkage also predicts a particular pattern of correlations among family members, viz., r (father-son) < r (mother- daughter) < r (mother-son) = r (father-daughter). Stafford (1961), a former employee of the Johnson O'Connor Research Foundation, was the first to publish family data in support of this hypothesis. Other studies with positive results soon appeared in the literature (e.g., Hartlage, 1970; Bock & Kolakowski, 1973; Yen, 1975). More recently, studies with larger samples (see Vandenberg & Kuse, 1979, for a review) have failed to replicate the predicted pattern of correlations among family members.

Bock, Zimowski, and Laciny (1986) recently compiled the parent-child correlations from studies that used relatively pure measures of spatial ability. Correlations from nine published studies that used the Identical Blocks

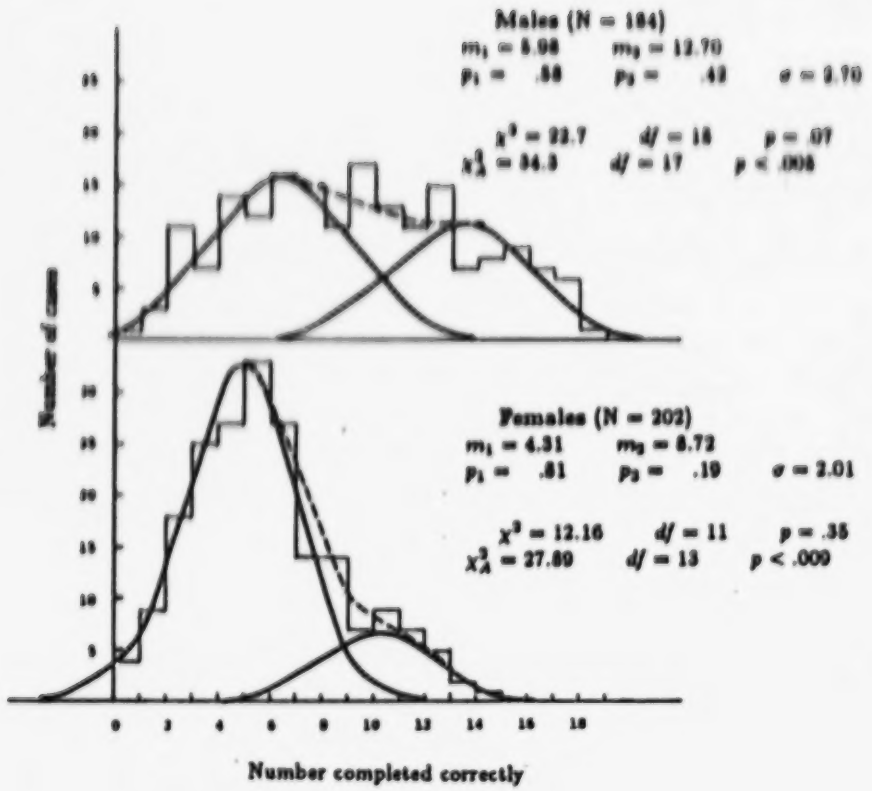


Figure 4: Score distributions by sex for the Vandenberg-Shepard test. Both distributions can be represented as the sum of two normal components. A single component does not fit the data. Adapted from Zimowski (1985).

test, the Mental Rotations test, or the Guilford-Zimmerman Spatial Visualization test qualified for inclusion. In cases where more than one analog measure was used in the same study, the correlations were averaged across measures.

The result of the analysis is shown in Figure 5. Even when only relatively pure measures of analog processing ability are considered, the accumulated evidence does not support the theory of recessive X-linkage. As the figure shows, the medians of the correlations do not follow the predicted pattern. Therefore, the inconsistencies among studies of recessive X-linkage do not appear due to impure measurement of the type we have described.

Recommendations For an Evaluation of the Measures of Structural Visualization Currently in Use at Johnson O'Connor

The test battery of the Johnson O'Connor Research Foundation currently includes two measures of Structural Visualization (spatial ability): the Paper-Folding test and the Wiggly Block test. The score reported for structural visualization is based on the average value of performance on both tests. In other words, these tests are treated as parallel (or congeneric) forms in the Foundation's assessment scheme (see Statistical Bulletin 1985-6).

Implicit in this treatment of the tests as parallel forms is the assumption that they not only measure a common spatial component or components but do so equally well. Available data, however, suggests that this may not be the case. Scores on the Paper-Folding and Wiggly Block tests correlate .39-.47 and .35-.38, respectively, with the analytical reasoning score of the battery and intercorrelate to an only slightly higher degree (.51-.65) (Technical Report 1983-2). It is thus uncertain whether these tests even measure a common spatial component: their intercorrelation could be entirely due to a shared analytic component (see Statistical Bulletin 1985-6).

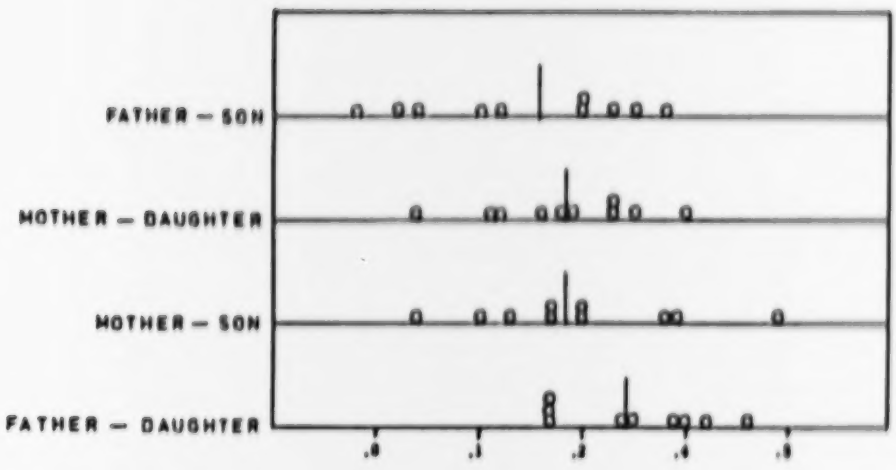


Figure 5: Distribution of familial correlations for analog spatial measures (after Bock, Zimowski, and Laciny, 1986).

In the remainder of this report we offer strategies for examining this measurement problem in greater detail. Our recommendations are motivated by the conclusions we have drawn in our examination of the literature and, in our judgment, represent the most potentially effectiv mean fo improving the Foundation's assessment of spatial ability.

First, we suggest a factor analysis of the Paper-Folding test at the *item level* to determine if it measures more than one ability. For these purposes, we recommend the full-information method for dichotomous items recently developed by Bock and Aitkin (1981) which provides a chi-square test for dimensionality. This factor analytic method has been successfully applied to other tests (see Bock, Gibbons, & Muraki, 1985; Zimowski, 1985), often with surprising results. Many cognitive tests thought to measure a single ability have been shown to measure more than one ability. Examinations of the pattern of item loadings has led to meaningful cognitive interpretations with implications for improving measurement (e.g., Zimowski, 1985).

Second, for comparative purposes, we recommend the administration of relatively pure measures of analog and nonanalog processing abilities along with the standard tests. For these purposes, we suggest three instruments. The first is a relatively pure measure of analog processing ability: a tape-cued, timed version of the Guilford-Zimmerman Spatial Visualization test developed by Bock and Kolakowski (1973). The second is the Incomplete Open Cubes test (Zimowski, 1985), which was especially constructed to measure analog and nonanalog processing in discrete item subsets. The third is a nonverbal reasoning test called Raven's Progressive Matrices, which was designed to measure Spearman's G in a culture-free manner. These tests have been described in greater detail elsewhere in this report and also in Statistical Bulletin 1985-16. Their inclusion in the evaluation study would provide supplemental information about the dimensional structure of the standard tests.

APPENDIX

Types of Spatial Test Items

This list provides examples of (mostly) paper-and-pencil item types that, at one time or another, have been called spatial items. All that these items in fact have in common is that the problems are presented in pictorial format.

Psychological test items are generally confidential. This protection is understandable from the test developer's view, but it makes it difficult to communicate test reviews to the professional community. We felt compelled to display only those items that have already been published elsewhere or items that are used for test instruction but are not typically scored. As a result, there is considerable variation in the difficulty of the items in this presentation. None of the most difficult items from any test are shown. Neither is the difficulty of the depicted items related to the difficulty of the entire test. The item presented from Raven's Progressive Matrices, for example, is an instruction probe that many children can understand, while the drawing of a Wiggly Block appears far more difficult than the actual presentation of the wooden Wiggly Block would be.

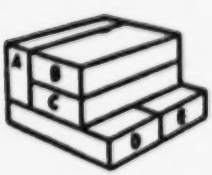
Not all the items presented in this section have been assembled into commercially available tests. Some types of items have been designed to provide data for very specific theoretical questions and have, so far, been used only in experimental laboratory settings. This is the case for the Attneave figures and the Palmer-Pylyshyn figures. Other items, like the Shepard-Metzler figures, were at first used exclusively in the laboratory, too, but have since appeared in paper-and-pencil form suitable for administration outside the laboratory.

- **Attneave Figures (Attneave & Arnoult, 1956), a.k.a. Random Shapes (Cooper, 1975)**

  The task is to determine whether both figures of a pair are rotated in the plane or whether one is also turned over. (Variations: different figures, affine contractions, translations, size scalings).

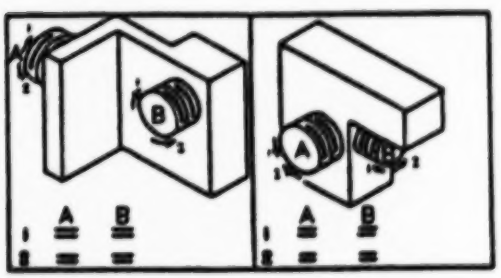
- **Block Counting (Thurstone, 1938)**

  A picture of a pile of blocks is shown. All blocks are the same size and shape. Some of them are marked with a letter. For each marked block, respondents have to determine how many other blocks it touches. The answers are written in the column on the right.



- **Bolts (Thurstone, 1951)**

  The items are pictures of blocks of wood. Two or three bolts are partially screwed into the wood. All the bolts have right-handed threads. The testee is required to indicate in which direction each bolt should be turned to drive it further into the wood.

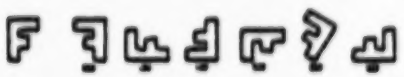


- **Camouflaged Outlines (Guilford & Lacey, 1947)**

  See Gottschaldt Figures.

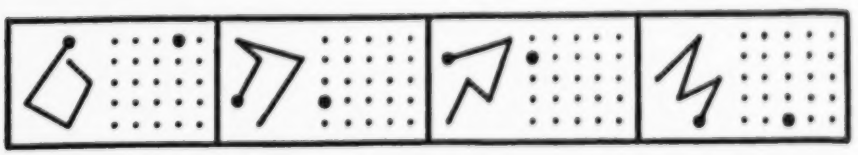
- **Cards (Thurstone & Thurstone, 1941)**

  Respondents are asked to mark the figures on the right that are rotated versions of the standard on the left. If a figure on the right is also turned over, it should not be marked.



- **Copying (Thurstone, 1938)**

  Respondents are asked to draw the figure on the left in the dotted space on the right. The little circles show where to begin. There is a dot for every corner.

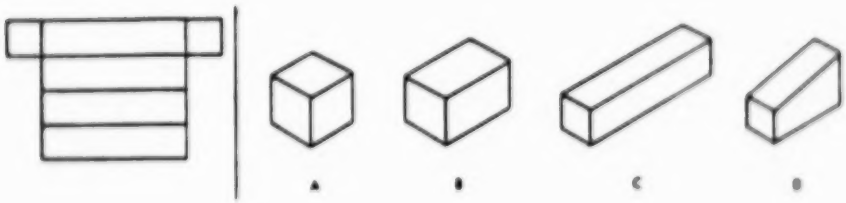


- **Cubes (Thurstone, 1938)**

  Pairs of cubes are shown. The respondents are told that each cube has six different faces. The task is to decide whether the two cubes in a pair *can* be the same or not.

- **DAT Space Relations (Bennett, Seashore, & Wesman, 1974)**

  The task is to decide which object on the right is the folded form of the unfolded surface on the left.

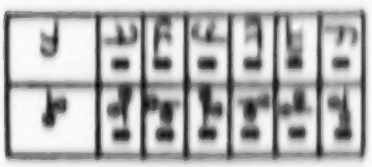


- **Embedded Figures (Witkin, 1950)**

  See Gottschaldt Figures.

- **Figures (Thurstone & Thurstone, 1941)**

  The task is to "mark every figure [in a row] which is like the first [one]. ... Figures which are made backward" are not to be marked.



- **Flags (Thurstone & Thurstone, 1941)**

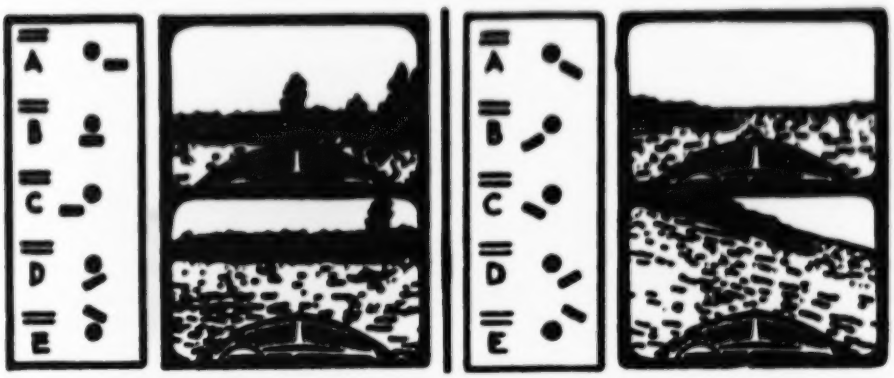
  Respondents are asked "to fit the pictures together by [mentally] sliding them around flat on the paper." Items are marked with a "+" sign if the two pictures show the same side of the flag, with a "–" sign if they show opposite sides.

- **Gottschaldt Figures (from Thurstone, 1944), a.k.a. Camouflaged Outlines (Guilford & Lacey, 1947), a.k.a. Embedded Figures (Witkin, 1950)**

  Respondents are asked to decide whether the simple figure shown on the left is contained in the more complex figures on the right. Each complex figure has to be evaluated.

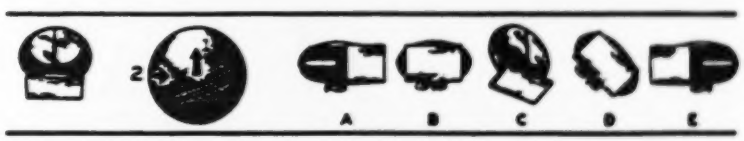
- **Guilford-Zimmerman Spatial Orientation Test (Guilford & Zimmerman, 1943)**

  Each item shows two views from the bow of a motorboat. Given the first view, respondents have to mark the change in orientation that yields the second view. The prow of the boat can move in three different ways: right-left, up-down, and clockwise vs. counterclockwise tilt.



- **Guilford-Zimmerman Spatial Visualization Test (Guilford & Zimmerman, 1947)**

  The test items present a picture of an alarm clock and instructions to mentally rotate the clock in specified directions. The task is to select the response alternative that shows the clock in its final position.



- **Hands (Thurstone, 1938)**

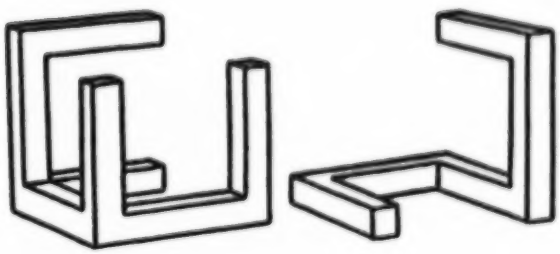
  Hands in different positions and orientations are presented. The task is to decide whether each picture shows the right or the left hand.

- **Identical Blocks (Educational Testing Service Staff, ca. 1951)**

  The items are pictures of blocks turned in different ways. The task is to decide which of the five blocks on the right is the same as the reference block on the left.

- **Incomplete Open Cubes (Zimowski, 1985)**

  The items in this test show two incomplete (i.e., parts of) open cubes. Testees are asked to determine whether the two incomplete cubes fit together to form a complete cube.

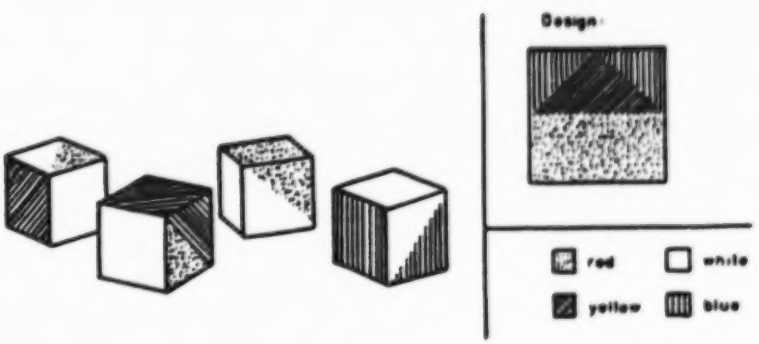


- **Kit Paper Form Board (French, Ekstrom, & Price, 1963), see also Minnesota Paper Form Board and Paper Form Board**

  Respondents have to decide which of five shaded forms on the right can be used to construct the unshaded figure on the left and which ones cannot. The decision is indicated with "+" and "–" signs.

- **Kohs' Block Design Test (Kohs, 1923)**

  Wooden cubes, like the ones on the left, are provided. The surfaces show several colors. Respondents are asked to arrange the cubes next to each other so that the design on the right is shown on the top surface.

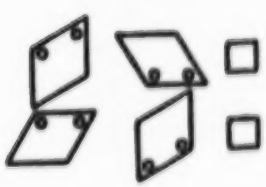


- **Lozenges A (Thurstone, 1938)**

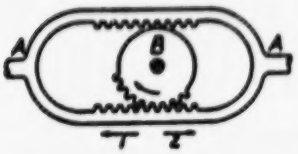
  The left figure "represents a lozenge-shaped card. It has a hole in one corner. It is painted black on one edge." The testee is told: "imagine that it is picked up, turned over, and placed *face down* so that the black edge touches the long black line in the figures on the right. Decide which of the two diagrams would fit. Where would the hole be?"

- **Lozenges B (Thurstone, 1938)**

  "Each pair of diagrams represents a lozenge-shaped card. If the two diagrams show the same side of the card, a plus sign is written in the square. If they show opposite sides of the card, a minus sign is written in the square."



- Mechanical Movements (Thurstone, 1938)



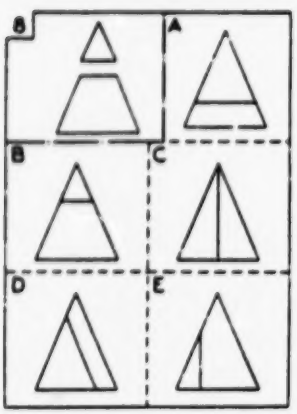
Respondents are asked several comprehension questions about the mechanical diagram:

If *B* starts moving in the direction shown, which way will *A* move, *1* or *2*?..... _____

In which direction will *A* be moving when *B* has turned half away around from where it is now? .... _____

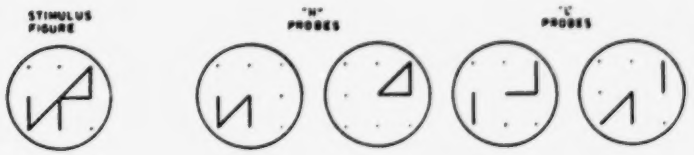
- Minnesota Paper Form Board (Likert, 1934; Likert & Quasha, 1970), see also Kit Paper Form Board and Paper Form Board

The upper left hand corner shows the parts of a geometric figure. Five response alternatives show the figure assembled from different sets of elements. Respondents are to mark the alternative that contains the pieces of the reference figure.



- **Palmer-Pylyshyn Figures (Palmer, 1978; Pylyshyn, 1979)**

  The participants are asked to decide whether the probes on the right are part of the reference figure on the left. "H" probes are "good" sub-figures under Gestalt principles, "L" probes are "poor" ones. Probes are presented in different orientations.

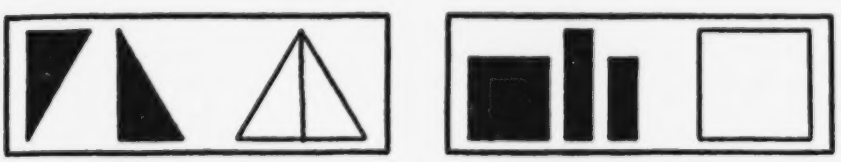


- **Paper Folding**

  See Punched Holes.

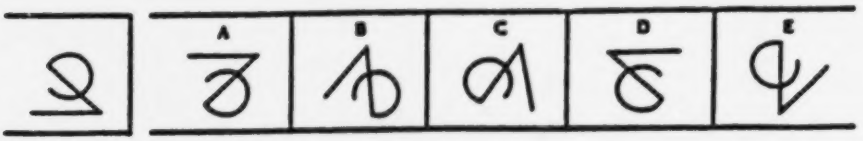
- **Paper Form Board (Thurstone, 1938), see also Minnesota Paper Form Board and KIT Paper Form Board**

  The black pieces on the left can be placed together to form the outline at the right. The task is to draw pencil lines in the white outlines to show how the black pieces may be placed to *fit* the outlines.



- PMA Space Test (Thurstone, 1958)

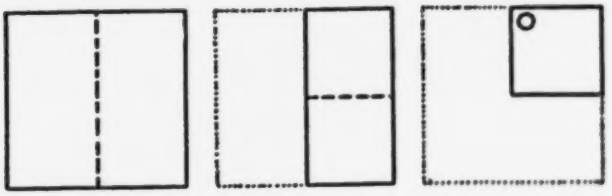
  Testees are asked to mark the probes that are rotated versions of the reference stimulus on the left. Probes that have also been turned over are not to be marked.



- Punched Holes (Thurstone, 1938), a.k.a. Paper Folding

  The left figure is a piece of paper that is to be folded along the dashed line. Then it will look like the middle figure. It is folded once more and then looks like the right figure. One or more holes are then punched through it.

  Respondents are asked to draw small circles on the first square to indicate where the holes will be when the paper is unfolded.



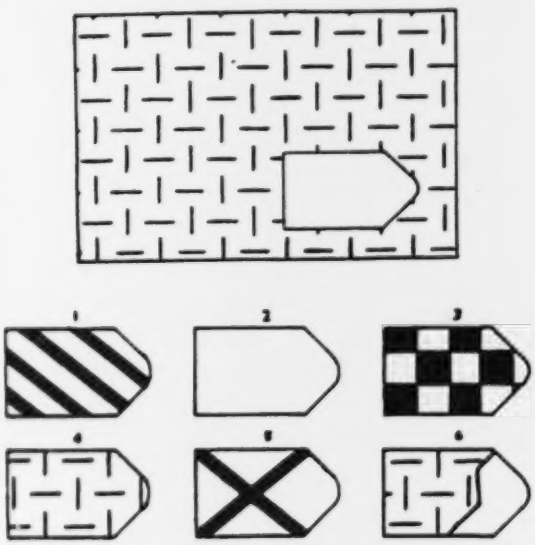
- Random Shapes (Cooper, 1975)

  See Attneave Figures.

- Raven's Progressive Matrices (Raven, 1938)

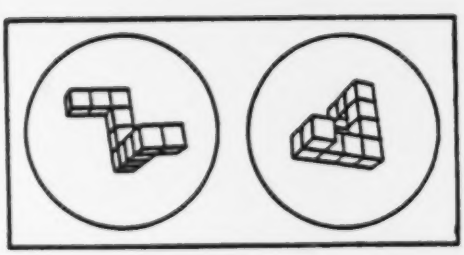
  Each item shows a matrix with a regularly constructed pattern. On the lower right of the matrix, a piece is missing.

  Respondents have to find the rules by which the pattern is constructed and identify the missing piece from six or eight response alternatives.



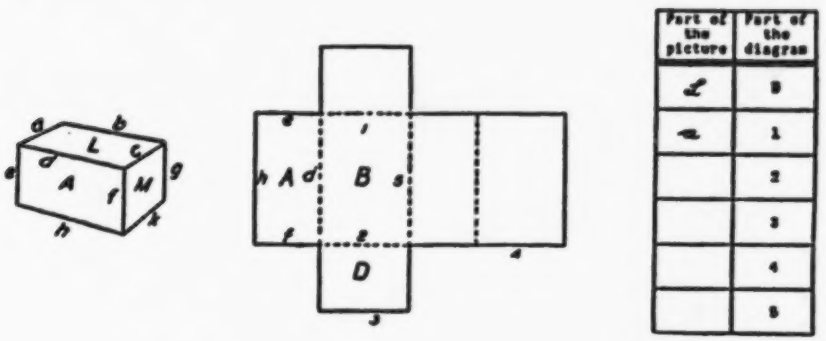
- Shepard-Metzler Block Figures (from Shepard, 1984)

  Decide whether the pairs of drawings show the same block figure in different orientations or whether different figures are shown.



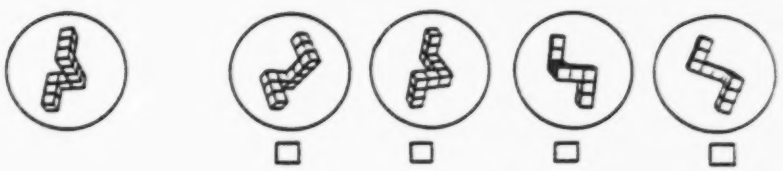
- **Surface Development (Thurstone, 1938)**

  A drawing of a three-dimensional object and its surface diagram are shown. Respondents have to match the corresponding elements of the figures.



- **Vandenberg-Shepard Mental Rotations Test (Vandenberg, 1971)**

  Respondents are asked to identify which two figures among the four shown on the right are rotated versions of the reference figure on the left.



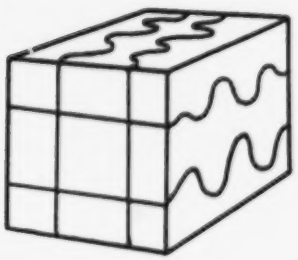
- **Wechsler's (WISC, WAIS) Block Design Test**

  A close relative of Kohs' Block Design test. The blocks show only two colors, white and red. Otherwise, the items have the same construction principles as Kohs'.

- Wiggly Block (O'Connor, 1928)

  A wooden block has been sawn into several pieces. The cuts are wavy or wiggly, rather than plane. All the pieces are approximately equal in volume and weight, but no two are identical.

  Respondents are presented with a random arrangement of the pieces and are asked to assemble the block as fast as possible.



References

Alderton, D.L., & Pellegrino, J.W. (1984). *Analysis of mental paper-folding.* Unpublished manuscript, University of California, Santa Barbara.

Anastasi, A. (1958). *Differential psychology: Individual and group differences in behavior* (3rd edition). New York: Macmillan.

Anderson, J.R. (1976). *Language, memory and thought.* Hillsdale, NJ: Erlbaum.

Anderson, J.R. (1978). Arguments concerning representations for mental imagery. *Psychological Review*, 34, 42-54.

Anderson, J.R. (1979). Further arguments concerning representations for mental imagery. *Psychological Review*, 86, 395-406.

Anderson, J.R. (1980). *Cognitive psychology and its implications.* San Francisco: Freedman, 1980.

Attneave, F., & Arnoult, M.D. (1956). The quantitative study of shape and pattern perception. *Psychological Bulletin*, 53, 221-227.

Baylor, G.W., & Racine, B. (1977) Mental imagery and the problems of cognitive representation: A computer approach. In J.M. Nicholas (Ed.), *Images, perception and knowledge.* Dordrecht, Holland: Reidel.

Bennett, G.K., Seashore, H.G., & Wesman, A. G. (1974). *Fifth edition manual for the differential aptitude tests forms S and T.* New York: The Psychological Corporation.

Bock, R.D. (1973). Word and image: Sources of the verbal and spatial factors in mental test scores. *Psychometrika*, 38, 437-457.

Bock, R.D., & Aitkin, M. (1981). Marginal Maximum Likelihood estimation of item parameters: An application of the EM algorithm. *Psychometrika*, 46, 443-459.

Bock, R.D., Gibbons, R.D., & Muraki, E. (1985). *Full-information factor analysis.* Chicago: National Opinion Research Center.

Bock, R.D., & Kolakowski, D. (1973). Further evidence of sex-linked major-gene influence on human spatial visualizing ability. *The American Journal of Human Genetics*, 25, 1-14.

Bock, R.D., Zimowski, M.F., & Laciny, C. (1986). *Sex differences in cognitive profiles: A developmental theory.* Submitted for publication.

Borich, G.D., & Bauman, P.M. (1972). Convergent and discriminant validation of the French and Guilford-Zimmerman spatial orientation and spatial visualization factors. *Educational and Psychological Measurement*, **32**, 1029-1033.

Brush, E.N. (1941). Mechanical ability as a factor in engineering aptitude. *Journal of Applied Psychology*, **25**, 300-312.

Bundesen, C., & Larsen, A. (1975). Visual transformation of size. *Journal of Experimental Psychology: Human Perception and Performance*, **1**, 214-220.

Bundesen, C., Larsen, A., & Farrell, J.E. (1982). Visual transformation of size and spatial orientation: Spatial imagery and perceptual processess. In A. Baddeley & J. Long (Eds.), *Attention and performance* (IX). Hillsdale, NJ: Erlbaum.

Burt, C. (1909). Experimental tests of general intelligence. *British Journal of Psychology*, **3**, 94-177.

Caplan, P.J., MacPherson, G.M., & Tobin, P. (1985). Do sex-related differences in spatial abilities exist?: A multilevel critique with new data. *American Psychologist*, **40**, 786-799.

Carroll, J.B. (1983). Studying individual differences in cognitive abilities: Through and beyond factor analysis. In R.F. Dillion & R.R. Schmeck (Eds.), *Individual differences in cognition* (Vol. 1, pp. 1-34). New York: Academic Press.

Connor, J.M., Schackman, M., & Serbin, L.A. (1978). Sex-related differences in response to practice in a visual-spatial test and generalized to a related test. *Child Development*, **49**, 24-29.

Cooper, L.A. (1975). Mental rotation of random two-dimensional shapes. *Cognitive Psychology*, **7**, 20-43.

Cooper, L.A. (1976a). Demonstration of a mental analog of an external rotation. *Perception and Psychophysics*, **19**, 296-302.

Cooper, L.A. (1976b). Individual differences in visual comparison processes. *Perception and Psychophysics*, **19**, 433-444.

Cooper, L.A. (1980a). Spatial information processing: Strategies for research. In R. Snow, P.A. Federico, & W.E. Montague (Eds.), *Aptitude, learning, and instruction: Cognitive process analysis.* Hillsdale, NJ: Erlbaum.

Cooper, L.A. (1980b). Recent themes in visual information processing: A selected overview. In R.E. Nickerson (Ed.), *Attention and performance* (Vol. 8). Hillsdale, NJ: Erlbaum.

Cooper, L.A., & Farrell, J.E. (1982). Effects of stimulus structure on rate of mental rotation. Unpublished manuscript.

Cooper, L.A., & Podgorny, P. (1976). Mental transformations and visual processes: Effects of complexity and similarity. *Journal of Experimental Psychology: Human Perception and Performance*, **2**, 503-514.

Cooper, L.A., & Shepard, R.N. (1973a). The time required to prepare for a rotated stimulus. *Memory and Cognition*, **1**, 246-250.

Cooper, L.A., & Shepard, R.N. (1973b). Chronometric studies of rotation of mental images. In W.G. Chase (Ed.), *Visual information processing.* New York: Academic Press.

Cooper, L.A., & Shepard, R.N. (1978). Transformations on representations of objects in space. In E.C. Carterette & M.P. Friedman (Eds.), *Handbook of Perception* (Vol. 8). New York: Academic Press.

Cunningham, J.P., Cooper, L.A., & Reaves, C.C. (1982). Visual comparison processes: Identity and similarity processes. *Perception and Psychophysics*, **32**, 50-60.

Daniel, M. (1978). *Wiggly Block worksample 3,4,5,6,7 manual.* New York: Johnson O'Connor Research Foundation.

Educational Testing Service Staff (ca. 1951). *Identical Blocks.* New York: College Entrance Examination Board.

Egan, D.E. (1978). *Characterizing spatial ability: Different mental processes reflected in accuracy and latency scores.* Pensacola, FL: Naval Aerospace Medical Research Laboratory.

Egan, D.E. (1979). Testing based on understanding: Implications from studies of spatial ability. *Intelligence*, **3**, 1-15.

Eliot, J.C., & Smith, I.M. (1983). *An international directory of spatial tests.* Windsor, England: NFER-Nelson.

French, J.W., Ekstrom, R.B., & Price, L.A. (1963). *Kit of reference tests for cognitive factors.* Princeton, NJ: Educational Testing Service.

French, J.W. (1951). The description of aptitude and achievement tests in terms of rotated factors. *Psychometric Monographs* (No. 5).

Funt, B.V. (1983). A parallel-process model of mental rotation. *Cognitive Science*, 7, 67-93.

Galton, F. (1883). *Inquiries into human faculty and its development.* London: Macmillan.

Guilford, J.P., & Lacey, J.I. [Eds.] (1947). *Printed classification tests.* Army Air Forces Aviation Psychology Program Research Report No. 5. Washington, DC: U.S. Government Printing Office.

Guilford, J.P., & Zimmerman, W.S. (1943). *The Guilford-Zimmerman aptitude survey – spatial orientation.* Beverly Hills, CA: Sheridan Supply Company.

Guilford, J.P., & Zimmerman, W.S. (1947). *The Guilford-Zimmerman aptitude survey – spatial visualization, form B.* Beverly Hills, CA: Sheridan Supply Company.

Hartlage, L.C. (1970). Sex-linked inheritance of spatial ability. *Perceptual and Motor Skills*, 31, 610.

Holliday, F. (1940). An investigation into selection of apprentices for the engineering laboratory. *Occupational Psychology*, 14, 69-81.

Holliday, F. (1943). The relations between psychological test scores and subsequent proficiency of apprentices in the engineering industry. *Occupational Psychology*, 17, 168-185.

Horn, J.L., & Cattell, R.B. (1966). Refinement and test of the theory of fluid and crystallized general intelligences. *Journal of Educational Psychology*, 57, 253-270.

Hunter, R.J. (1945). Aptitude tests for the machine shop. *Industrial Arts and Vocational Education*, 34, 58-64.

Hyde, J.S. (1981). How large are cognitive gender differences? A meta- analysis using w and d. *American Psychologist*, 36, 892-901.

Kelley, T.L (1928). *Crossroads in the mind of man.* Stanford, CA: Stanford University Press.

Kohs, S.C. (1923). *Intelligence measurement.* New York: Macmillan.

Kosslyn, S.M. (1980). *Image and mind.* Cambridge, MA: Harvard University Press.

Kosslyn, S.M. (1981). The medium and the message in mental imagery: A theory. *Psychological Review*, 88, 46-66.

Kosslyn, S. M., & Shwartz, S.P. (1977). A simulation of visual imagery. *Cognitive Science*, 1, 265-295.

Kubovy, M., & Podgorny, P. (1981). Does pattern matching require the normalization of size and orientation? *Perception and Psychophysics*, 30, 24-28.

Kyllonen, P.C., Lohman, D.F., & Snow, R.E. (1984). Effects of aptitudes, strategy training, and task facets on spatial task performance. *Journal of Educational Psychology*, 76, 130-145.

Lansman, M. (1981). Ability factors and the speed of information processing. In M.P. Friedman, J.P. Das, & N. O'Connor (Eds.), *Intelligence and learning* (pp.441-447). New York: Plenum Press.

Larsen, A., & Bundesen, C. (1978). Size scaling in visual pattern recognition. *Journal of Experimental Psychology: Human Perception and Performance*, 4, 1-20.

Levine, S., & Huttenlocher, J. (1985). Sex differences in spatial ability of preschool children. Unpublished manuscript, The University of Chicago, Chicago, IL.

Likert, R. (1934). A multiple-choice revision of the Minnesota Paper Form Board. *Psychological Bulletin*, 31, 674.

Likert, R., & Quasha, W.H. (1970). *Manual for the revised Minnesota Paper Form Board test.* New York: The Psychological Corporation.

Linn, M.C., & Petersen, A.C. (1985). Emergence and characterization of sex differences in spatial ability: A meta-analysis. *Child Development*, 56, 1479-1498.

Lohman, D.F. (1979). *Spatial ability: individual differences in speed and level* (Technical Report No. 9). Stanford, CA: Aptitude Research Project, School of Education, Stanford University.

Lohman, D.F., & Kyllonen, P.C. (1983). Individual differences in solution strategy on spatial tasks. In R.F. Dillion & R.R. Schmeck (Eds.), *Individual differences in cognition* (Vol 1, pp. 105-135). New York: Academic Press.

Maccoby, E.E., & Jacklin, C.N. (1974). *The psychology of sex differences.* Stanford,CA: Stanford University Press.

Martin, G.C. (1951). Test batteries for automechanics and apparel design. *Journal of Applied Psychology*, 35, 20-22.

McClearn, G. E., & DeFries, J.C. (1973). *Introduction to behavioral genetics.* San Francisco: W.H. Freeman.

McFarlane, M. (1925). A study of practical ability. *British Journal of Psychology, Monograph Supplement No. 8.*

McGee, M.G. (1979). *Human spatial abilities.* New York: Holt, Rinehart, and Winston.

Metzler, J. (1973). *Cognitive analogies of the rotation of three-dimensional objects.* Unpublished doctoral dissertation, Stanford University, Stanford, CA.

Metzler, J., & Shepard, R.N. (1974). Transformational studies of the internal representation of three-dimensional structures. In Solso, R.L. (Ed.) *Theories in cognitive psychology: The Loyola symposium* (pp.147- 201). New York: Wiley.

Michael, W.B., Guilford, J.P., Fruchter, B., & Zimmerman, W.S. (1957). The description of spatial-visualization abilities. *Educational and Psychological Measurement*, **17**, 185-199.

Moran, T.P. (1973). *The symbolic nature of visual imagery.* Paper presented at the third International Conference on Artificial Intelligence, Stanford, CA.

Mumaw, R.J., & Pellegrino, J.W. (1984). Individual differences in complex spatial processing. *Journal of Educational Psychology*, **76**, 920-939.

Mumaw, R.J., Pellegrino, J.W., & Glaser, R. (1980, November). *Some puzzling aspects of spatial ability.* Paper presented at the annual meeting of Psychonomic Society, St. Louis.

Mumaw, R.J., Pellegrino, J.W., Kail, R.V., Jr., & Carter, P. (1984). Different slopes for different folks: Process analysis of spatial aptitude. *Memory and Cognition*, **12**, 515-521.

Nash, S.C. (1975). The relationship among sex-role stereotyping, sex- role preference, and the sex difference in spatial visualization. *Sex Roles*, **1**, 15-32.

O'Connor, J. (1928). *Born that way.* Baltimore: Williams & Wilkins.

O'Connor, J. (1943). *Structural visualization.* Boston: Human Engineering Laboratory.

Paivio, A. (1969). Mental imagery in association learning and memory. *Psychological Review*, **3**, 241-263.

Paivio, A. (1977). Images, propositions, and knowledge. In J.M. Nicholas (Ed.), *Images, perception and knowledge.* Dordrecht, Holland: Reidel.

Palmer, S.E. (1977). Hierachical structure in perceptual representation. *Cognitive Psychology*, **9**, 441-474.

Palmer, S.E. (1978). Fundamental aspects of cognitive representation. In E.H. Rosch & B.B. Lloyd (Eds.), *Cognition and categorization.* Hillsdale, NJ: Erlbaum.

Pellegrino, J.W., Alderton, D.L., & Regian, J.W. (1984, December). *Components of spatial ability.* Paper presented at the NATO Advanced Study Institute in Cognition and Motivation, Athens, Greece.

Pellegrino, J.W., Alderton, D.L., & Shute, V.J. (1984). Understanding spatial ability. *Educational Psychologist*, **19**, 239-253.

Pellegrino, J.W., Cantoni, V.J., & Solter, A. (1981, November). *Speed, accuracy, and strategy differences in spatial processing.* Paper presented at the annual meeting of the Psychonomic Society, Philadelphia.

Pellegrino, J.W., & Glaser, R. (1979). Cognitive correlates and components in the analysis of individual differences. *Intelligence*, **3**, 187-214.

Pellegrino, J.W., & Goldman, S. R. (1983). Developmental and individual differences in verbal and spatial reasoning. In R.F. Dillon & R.R. Schmeck (Eds.), *Individual differences in cognition.* (Vol. 1). New York: Academic Press.

Pellegrino, J.W., & Mumaw, R.J. (1980). Multicomponent models of spatial ability. Unpublished manuscript, University of California, Santa Barbara, CA.

Pellegrino, J.W., Mumaw, R.J., & Shute, V.S. (1985). Analysis of spatial aptitude and expertise. In S.E. Embretson (Ed.), *Test design: Contributions from psychology, education and psychometrics.* New York: Academic Press.

Petersen, A.C. (1976). Physical androgyny and cognitive functioning in adolescence. *Developmental Psychology*, **12**, 524-533.

Petersen A.C., Tobin-Richards, M.H., & Crockett, L. (1982). Sex differences. In H.E. Mitzel (Ed.), *Encyclopedia of Educational Research* (Fifth Edition). New York: The Free Press.

Plant, W.T., Southern, M.L., & Jacklin, C.N. (1977). *Statistically significant sex differences in attitude, interest, and personality measures: Much ado about very little!* Paper presented at the meeting of the Western Association for Women in Psychology, San Jose, CA.

Podgorny, P. (1975). Mental rotation and the third dimension. Unpublished senior honors thesis, Stanford University, Stanford, CA.

Price, L., & Eliot, J. (1975). Convergent and discriminant validities of two sets of measures of spatial orientation and visualization. *Educational and Psychological Measurement*, **35**, 975-977.

Pylyshyn, Z.W. (1979). The rate of mental rotation of images: A test of a holistic analogue hypothesis. *Memory and Cognition*, **7**, 19-28.

Raven, J.C. (1938). *Progressive Matrices.* Cambridge, England: Cambridge University Press.

Raven, J.C. (1962). *Advanced Progessive Matrices, Set II* (1962 Revision). London: E.T. Heron & Co.

Regard, M., Strauss, E., & Knapp, P. (1982). Children's production on verbal and non-verbal fluency tasks. *Perceptual and Motor Skills*, **55**, 839-844.

Remmers, H.H., & Schell, J.W. (1933). Testing the O'Connor Wiggly Block test. *Personnel Journal*, **12**, 155-159.

Richmond, P.G. (1980). A limited sex difference in spatial test scores with a preadolescent sample. *Child Development*, **53**, 1106-1110.

Roff, M. (1952). A factorial study of tests in the perceptual area. *Psychometric Monographs No. 8.*

Rosser, R.A., Ensing, S.S., Glider, P.J., & Lane, S. (1984). An information-processing analysis of children's accuracy in predicting the appearance of rotated stimuli. *Child Development*, **55**, 2204-2211.

Sekuler, R., & Nash, D. (1972). Speed of size scaling in human vision. *Psychonomic Science*, **27**, 93-94.

Shepard, R.N. (1978). The mental image. *American Psychologist*, **33**, 125-137.

Shepard, R.N. (1981). Psychophysical complementarity. In M. Kubovy and J.R. Pomerantz (Eds.), *Perceptual organization.* Hillsdale, NJ: Erlbaum.

Shepard, R.N. (1984). Ecological constraints on internal representation: Resonant kinematics of perceiving, imagining, thinking, and dreaming. *Psychological Review*, **91**, 417-447.

Shepard, R.N., & Cooper, L.A. (1982). *Mental images and their transformations.* Cambridge, MA: MIT Press.

Shepard, R.N., & Feng, C. (1972). A chronometric study of mental paper folding. *Cognitive Psychology*, **3**, 228-243.

Shepard, R.N., & Metzler, J. (1971). Mental rotation of three- dimensional objects. *Science*, **171**, 701-703.

Sherman, J. (1980). Mathematics, spatial visualization and related factors: Changes in boys and girls, grades 8-11. *Journal of Educational Psychology*, **73**, 476-482.

Slater, P. (1941). Tests for selecting secondary and technical school children. *Occupational Psychology*, **14**, 40-55.

Smith, I.M. (1964). *Spatial ability.* San Diego, CA: Robert R. Knapp.

Snow, R.E. (1978). Eye fixation and strategy analyses of individual differences in cognitive aptitudes. In A.M. Lesgold, J.W. Pellegrino, S.D. Fokkema, & R. Glaser (Eds.), *Cognitive psychology and instruction* (pp.229-308). New York: Plenum.

Snow, R.E. (1980). Aptitude processes. In R.E. Snow, P.A. Federico, & W.E. Montague (Eds.), *Aptitude, learning, and instruction* (Vol. 1). Hillsdale, NJ: Erlbaum.

Snow, R.E., Kyllonen, P.C., & Marshalek, B. (in press). The topography of ability and learning correlations. In R.J. Sternberg (Ed.), *Advances in the psychology of human intelligence* (Vol. 2). Hillsdale, NJ: Erlbaum.

Spearman, C. (1904). "General intelligence," objectively determined and measured. *American Journal of Psychology*, **15**, 201-293.

Spearman, C., & Jones, L.W. (1950). *Human ability.* London: Macmillan.

Stafford, R.E. (1961). Sex differences in spatial visualization as evidence of sex-linked inheritance. *Perceptual and Motor Skills*, **13**, 428.

Statistical Bulletin 438. (1940). *Validation study of elementary school teachers.* L. Seeley. Boston: Johnson O'Connor Research Foundation.

Statistical Bulletin 706. (1941). *Validation study of 35 magazine editorial staff members.* L. Seeley. Boston: Johnson O'Connor Research Foundation.

Statistical Bulletin 1985-6. *Research proposal: Analysis of structural visualization tests.* W. Wothke, M.F. Zimowski, & R. Smith. New York: Johnson O'Connor Research Foundation.

Statistical Bulletin 1985-16. *Research proposal: Description of experimental items of structural visualization.* W. Wothke & M.F. Zimowski. New York: Johnson O'Connor Research Foundation.

Sternberg, R.J. (1977). *Intelligence, information processing, and analogical reasoning: The componential analysis of human abilities.* Hillsdale, NJ: Erlbaum.

Strauch, A.B. (1976, September). *The development of sex differences in spatial and verbal abilities.* Paper presented at the meeting of the American Psychological Association, Washington, DC.

Tapley, S.M., & Bryden, M.P. (1977). An investigation of sex differences in spatial ability: Mental rotation of three-dimensional objects. *Canadian Journal of Psychology*, **31**, 122-130.

Taylor, L.J. (1977). Sex and psychological differentiation. *Psychological Reports*, **41**, 192-194.

Technical Report 81. (1940). *Analysis of the worksample scores of forty-nine accountants, Report 1.* L. Seeley. Boston: Johnson O'Connor Research Foundation.

Technical Report 90. (1941). *Analysis of the worksample results of 101 high school teachers.* L. Seeley. Boston: Johnson O'Connor Research Foundation.

Technical Report 97. (1941). *A follow-up of the careers of tested engineering students: Part I.* S. Horton. Boston: Johnson O'Connor Research Foundation.

Technical Report 97. (1942). *A follow-up of the careers of tested engineering students: Part II.* S. Horton. Boston: Johnson O'Connor Research Foundation.

Technical Report 102. (1942). *Preliminary study of the relationship of test scores to success and failure in school.* M. Luqueer. Boston: Johnson O'Connor Research Foundation.

Technical Report 113. (1943). *A study of the correlations between worksamples 1, 4, and 5, 16, 18, and 35 and subjects given a senior high school A.* K. Rippere. Boston: Johnson O'Connor Research Foundation.

Technical Report 1981-2. *Construct validity of the Wiggly Block and Paper Folding tests: A factor-referenced multivariate study.* R.S. Trafton. Boston: Johnson O'Connor Research Foundation.

Technical Report 1982-6. *A factorial study of reasoning tests.* M. Daniel. Boston: Johnson O'Connor Research Foundation.

Technical Report 1983-2. *Large-sample test intercorrelations.* M. Daniel. New York: Johnson O'Connor Research Foundation.

Thissen, D. (1982). Marginal Maximum Likelihood estimation for the one-parameter logistic model. *Psychometrika*, **47**, 175-186.

Thorndike, E.L. (1921). On the organization of intellect. *Psychological Review*, **28**, 141-151.

Thurstone, L.L. (1938). Primary mental abilities. *Psychometric Monographs No.1.*

Thurstone, L.L. (1944). *A factorial study of perception.* Chicago: The University of Chicago Press.

Thurstone, L.L. (1950). *Some primary abilities in visual thinking.* Chicago: The University of Chicago Psychometric Laboratory Report No. 59.

Thurstone, L.L. (1951). *An analysis of mechanical aptitude.* Chicago: The University of Chicago Psychometric Laboratory Report No. 62.

Thurstone, L.L., & Thurstone, T.G. (1941). Factorial studies of intelligence. *Psychometric Monographs No. 2.*

Thurstone, T.G. (1958). *Manual for the SRA Primary Mental Abilities.* Chicago: Science Research Associates.

Validation Bulletin 11. (1941). *Confidential report to company 13.* C. Wilks. Boston: Johnson O'Connor Research Foundation.

Validation Bulletin 22a. (1942). *Analysis of the worksample results of twenty aviation mechanics.* J. Ricks. Boston: Johnson O'Connor Research Foundation.

Validation Bulletin 74. (1947). *A study of company 11.* M.H. Licht. Boston: Johnson O'Connor Research Foundation.

Vandenberg, S.G. (1971). *The mental rotations test.* Boulder: University of Colorado.

Vandenberg, S.G. (1975). Sources of variance in performance of spatial tests. In J. Eliot & N.J. Salkind (Eds.), *Children's spatial development.* Springfield, IL: C.C. Thomas.

Vandenberg, S.G., & Kuse, A.R. (1979). Spatial ability: A critical review of the sex-linked major gene hypothesis. In M.A. Wittig & A.C. Petersen (Eds.), *Sex-related differences in cognitive functioning: Developmental issues.* New York: Academic Press.

Waber, D.P. (1977). Sex differences in mental abilities, hemispheric lateralization, and rate of physical growth at adolescence. *Developmental Psychology*, **13**, 29-38.

Witkin, H.A. (1950). Individual differences in ease of perception of embedded figures. *Journal of Personality*, **19**, 1-15.

Wolfe, D. (1940). Factor analysis to 1940. *Psychometric Monograph No. 3.*

Yen, W.M. (1975). Sex-linked major-gene influences on selected types of spatial performance. *Behavior Genetics*, **5**, 405-409.

Yuille, J.C., & Steiger, J.H. (1982). Nonholistic processing in mental rotation: Some suggestive evidence. *Perception and Psychophysics*, **31**, 201-209.

Zimowski, M.F. (1985). *Attributes of spatial test items that influence cognitive processing.* Unpublished doctoral dissertation, The University of Chicago, Department of Behavioral Sciences.

JOHNSON & JOHNSON RESEARCH FOUNDATION
HUMAN ENGINEERING LABORATORY